Naval Research Laboratory

Washington, DC 20375-5000

AD-A224 056

Wald Sequential Detection with Non-Gaussian Pulsed Radar Data Using the Zakai Equation

SERAFIN P. RODRIQUEZ

Advanced Technology Branch
Tactical Electronic Warfare Division

July 20, 1990

DTIC
ELECTE
JUL 18 1990
S   D
E

90 07 17 030

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1990 July 20	Final

4. TITLE AND SUBTITLE

Wald Sequential Detection with Non-Gaussian Pulsed Radar Data Using the Zakai Equation

5. FUNDING NUMBERS

PE - 63270N
PR - R2030E00T0
WU - 0N156-065

6. AUTHOR(S)

S. P. Rodriquez

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Maryland
College Park, MD 20740

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL Memorandum
Report 6686

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Technology
Arlington, VA 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The "optimal" Wald sequential hypothesis test for diffusion signals is presented. The result is a threshold test with explicitly computable thresholds. Five possible schemes for a numerical implementation of the test are given. A comparison of the different implementations and analysis of the detectors performance is done for the radar problem of ship versus chaff target discrimination using lognormal and Rayleigh models respectively. Parameter estimation for the lognormal and Rayleigh cases is also studied. Finally, a signal estimation scheme is presented utilizing the conditional expectation of the signal computed from the conditional density of the underlying state, which is the solution to the Zakai equation.

14. SUBJECT TERMS

Detection, Likelihood ratio detectors, Optimal discrimination, π, $\in$

15. NUMBER OF PAGES

93

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

CONTENTS

Chapter 1

Introduction

This thesis implements an "optimal" Wald sequential hypothesis testing scheme
for the general diffusion model,

$$dx_t = f^i(\mathbf{x}_t)dt + g^i(\mathbf{x}_t)d\mathbf{W}_t^i \qquad (1.1)$$

$$dy_t = h^i(\mathbf{x}_t)dt + \gamma d\mathbf{V}_t \qquad (1.2)$$

where $i = 0, 1$. Under hypothesis i;

$\mathbf{x}_t$ is an n^i dimensional state vector;

$\mathbf{W}_t^i$ is an m^i dimensional Brownian motion vector;

$\mathbf{y}_t$ is a p dimensional observation vector;

$\mathbf{V}_t$ is a p dimensional Brownian motion vector;

identical under both hypothesis;

f^i, g^i, h^i are known functions of the respective state vectors.

So $f^i(\mathbf{x}_t)$ is an n^i vector, $g^i(\mathbf{x}_t)$ is an $n^i \times m^i$ matrix, and $h^i(\mathbf{x}_t)$ is a p vector.
Equation 1.1 is called the state equation and equation 1.2 is the observation
equation. When necessary, we will drop the time subscript and use subscripts
to denote elements in vectors or matrices.

Given observation data from one of the hypotheses, we wish to determine
from which hypothesis the data came. Quoting results from [LS78] we will

present an "optimal" sequential detection scheme and then formulate a solution with several possible numerical implementations of the detector. We will give some numerical results and compare several of these numerical schemes through computer simulation.

To begin, let us review some results from chapter 17 of [LS78]. Consider the following pair of stochastic differential equations

$$
\begin{aligned}
H_0 \quad &: \quad dy_t = dV_t \\
H_1 \quad &: \quad dy_t = h(x_t)dt + dV_t \quad y_0 = 0
\end{aligned}
$$

These correspond to scalar observation equations for a single hypothesis versus noise case. Let $(\Omega, \mathcal{F}, P)$ be a given probability space with a nondecreasing family of σ-algebras $\mathcal{F}_t$ where $t \geq 0$ with $\mathcal{F}_t \subseteq \mathcal{F}$. Furthermore, let $W = (W_t, \mathcal{F}_t)$ be a Wiener process. Assume $h(x_t) = (h(x_t), \mathcal{F}_t)$ is an unobservable process independent of W. We make the following definitions which describe the sequential detection scheme.

Definition : A sequential hypothesis testing scheme is denoted by $\Delta = \Delta(\tau, \delta)$. $\tau(y)$ signifies the decision time and $\delta(y) \in \{0, 1\}$ the decision with $\delta = 1$ signifying the observed data came from the model corresponding to hypothesis 1 and $\delta = 0$ from hypothesis 0.

Definition : We define an error of the first kind as $\alpha(\Delta) = P_1(\delta(y) = 0)$ which in radar terminology is called the probability of miss and denoted by P_M. We define an error of the second kind as $\beta(\Delta) = P_0(\delta(y) = 1)$ which in radar terminology this is called the probability of false alarm and denoted by P_F.

Definition : We define $\Delta_{\alpha,\beta}$ as the class of schemes $\Delta = \Delta(\tau, \delta)$ with $\alpha(\Delta) \leq \alpha$ and $\beta(\Delta) \leq \beta$ where α, β are constants with $\alpha + \beta < 1$.

Definition : Define $\hat{h}_t = E_1(h(x_t)|\mathcal{F}_t^y)$

We make the following assumptions required for the proof of the following theorem,

A1 $E_i\left[|h(x_t)|\right] < \infty$ $t < \infty$, $i = 0, 1$

A2 $P_i\left\{\int_0^\infty \hat{h}_s^2 ds = \infty\right\} = 1$ $i = 0, 1$

A3 $E_i\left[\int_0^t \hat{h}_s^2 ds\right] < \infty$ $t < \infty$, $i = 0, 1$.

Theorem 1 *Given assumptions A1 to A3 then in the class $\Delta_{\alpha,\beta}$ there exists a scheme $\tilde{\Delta} = \Delta(\tilde{\tau}, \tilde{\delta})$ optimal in the sense that for any other scheme $\Delta = \Delta(\tau, \delta) \in \Delta_{\alpha,\beta}$*

$$E_i\left[\int_0^{\tilde{\tau}} \hat{h}_t^2 dt\right] \leq E_i\left[\int_0^\tau \hat{h}_t^2 dt\right]   i = 0, 1 .$$

The scheme $\tilde{\Delta} = \Delta(\tilde{\tau}, \tilde{\delta})$ can be defined by the relationship

$$\tilde{\tau}(y) = \text{infimum}\left\{t \geq 0 : \Lambda_t(y) \notin (A, B)\right\}$$

$$\tilde{\delta}(y) = \begin{cases} 0 & \Lambda_{\tilde{\tau}(y)}(y) \geq B \\ 1 & \Lambda_{\tilde{\tau}(y)}(y) \leq A \end{cases}$$

where

$$\Lambda_t(y) = \exp\left\{\int_0^t \hat{h}_s dy_s - \frac{1}{2}\int_0^t \hat{h}_s^2 ds\right\}$$

and

$$A = \frac{\alpha}{1 - \beta}   ,   B = \frac{1 - \alpha}{\beta}   .$$

Λ_t is known as the likelihood ratio and $\ln(\Lambda_t)$ is the log likelihood ratio. In this case

$$E_i\left[\int_0^{\tilde{\tau}} \hat{h}_t^2 dt\right] \leq \infty   i = 0, 1$$

Lemma 1 *For the scheme $\tilde{\Delta} = \Delta(\tilde{\tau}, \tilde{\delta})$,*

$$P_0(\tilde{\tau}(y) < \infty) = P_1(\tilde{\tau}(y) < \infty) = 1$$

and $\alpha(\tilde{\Delta}) = \alpha, \beta(\tilde{\Delta}) = \beta$.

The proofs for the above theorem and lemma are given in Chapter 17 of [LS78] for the scalar case.

The proofs for the equivalent theorem and lemma for the vector case are given in [LaV86]. For the vector case assumptions A1 through A3 become:

3

B1 $E_i\left[\|h(\mathbf{x}_t)\|\right] < \infty \quad t < \infty \, , \, i = 0, 1$

B2 $P_i\left\{\int_0^\infty \|\hat{h}_s\|^2 ds = \infty\right\} = 1 \quad i = 0, 1$

B3 $E_i\left[\int_0^t \|\hat{h}_s\|^2 ds\right] < \infty \quad t < \infty \, , \quad i = 0, 1$

and then

$$\Lambda_t(y) = \exp\left\{\int_0^t \hat{h}_s^T dy_s - \frac{1}{2}\int_0^t \|\hat{h}_s\|^2 ds\right\}$$

where ^{T} in $\hat{h}_s^T$ denotes vector transpose.

For the general hypothesis testing problem between two hypotheses, corresponding to eqs. 1.1 and 1.2, the likelihood ratio is given by

$$\Lambda_t(y) = \exp\left\{\int_0^t \left(\hat{h}_s^1 - \hat{h}_s^0\right)^T dy_s - \frac{1}{2}\int_0^t \left(\|\hat{h}_s^1\|^2 - \|\hat{h}_s^0\|^2\right) ds\right\} \quad .$$

With the additional assumption

$$P_i\left\{\int_0^\infty \|\hat{h}_s^1 - \hat{h}_s^0\|^2 ds = \infty\right\} = 1 \quad i = 0, 1$$

we guarantee a finite decision time.

In order to calculate $\Lambda_t(y)$ it is necessary to determine $\hat{h}_t^i$, the conditional expectation of $h^i(\mathbf{x}_t)$ given the observation y_s, $s \le t$. Under appropriate conditions, the conditional density of $\mathbf{x}_t$ conditioned on the observation dy_t, which we will denote as $u_i(\mathbf{x}, t)$, satisfies the linear parabolic partial differential equation known as the Zakai equation [BBH83]. In the next chapter we present the Zakai equation and the necessary conditions for existence of a unique solution which in turn will yield $\Lambda_t(y)$. We also derive several approximation schemes which are implemented and we make comparisons of the results.

Chapter 2

Solution and Numerical Approximation of the Zakai Equation

2.1 Solution of the Zakai equation

The Zakai equation is given by (2.1). When γ (the observation noise scaling factor) is equal to one the conditional density $u(\mathbf{x},t)$ of the underlying state $\mathbf{x}$ conditioned on the observations dy is known to satisfy the Zakai equation. We will drop the hypothesis superscripts and the time subscripts, instead using subscripts to denote vector and matrix elements. So the functions $h(\mathbf{x}_t)$, $\hat{h}_t$, $\mathbf{y}_t$, and $\mathbf{x}_t$ will not have explicit time dependence. The following assumptions guarantee existence and uniqueness of the solution to the Zakai equation [BBH83] and are assumed to hold throughout the chapter.

C1 L^* is uniformly elliptic

C2 $f(\mathbf{x}),\sigma(\mathbf{x}),h(\mathbf{x})$, $\frac{\partial}{\partial x_i}f_i(\mathbf{x}),\frac{\partial}{\partial x_i}\sigma_{i,j}(\mathbf{x}),\frac{\partial}{\partial x_j}\sigma_{i,j}(\mathbf{x}),\frac{\partial^2}{\partial x_i\partial x_j}\sigma_{i,j}(\mathbf{x}),\frac{\partial}{\partial x_i}h_k(\mathbf{x})$, and $\frac{\partial^2}{\partial x_i\partial x_j}h_k(\mathbf{x})$ for $i,j = 1,...,n$ and $k = 1,...,p$ are uniformly bounded and Lipschitz continuous.

Then the unormalized conditional density $u(\mathbf{x}, t)$ satisfies (2.1),

$$du(\mathbf{x},t) \;=\; L^* u(\mathbf{x},t)dt - \frac{1}{2}\|h\|^2 u(\mathbf{x},t)dt + h^T dy u(\mathbf{x},t) \qquad (2.1)$$

$$L^* u(\mathbf{x},t) = \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} \left[\sigma_{i,j}(\mathbf{x}) u(\mathbf{x},t) \right] - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left[f_i(\mathbf{x}) u(\mathbf{x},t) \right]$$

$$\sigma(\mathbf{x}) = \frac{1}{2} g(\mathbf{x}) g^T(\mathbf{x})$$

where $\sigma_{i,j}(\mathbf{x})$ is the i, j element of the the matrix $\sigma(\mathbf{x})$. To determine the partial differential equation that $u(\mathbf{x}, t)$ satisfies for $\gamma \neq 1$ it is simply neccesary to renormalize the output equation, so $h(\mathbf{x})$ becomes $\frac{1}{\gamma} h(\mathbf{x})$ and the observation dy becomes $\frac{1}{\gamma} dy$.

We wish to solve (2.1) for $u(\mathbf{x}, t)$. We factor the term $L^* u(\mathbf{x}, t)$ as $A^* u(\mathbf{x}, t) + C(\mathbf{x}) u(\mathbf{x}, t)$ where A^* contains the terms which have derivatives of the density $u(\mathbf{x}, t)$. This is done based on numerical implementation considerations which will be expanded upon later. Then (2.1) can be written as

$$du(\mathbf{x},t) \;=\; A^* u(\mathbf{x},t)dt + \left[C(\mathbf{x}) - \frac{1}{2}\|h\|^2 \right] u(\mathbf{x},t)dt + h^T dy u(\mathbf{x},t) \quad . \ (2.2)$$

Defining

$$d\phi(\mathbf{x},\mathbf{y},t) = h^T dy + \left[C(\mathbf{x}) - \frac{1}{2}\|h\|^2 \right] dt$$

we will use a Gauge Transformation with

$$r(\mathbf{x},t) = \exp^{-\phi(\mathbf{x},\mathbf{y},t)} u(\mathbf{x},t)$$

so

$$u(\mathbf{x},t) = \exp^{\phi(\mathbf{x},\mathbf{y},t)} r(\mathbf{x},t) \quad .$$

Now $u(\mathbf{x}, t)$ must satisfy (2.2) so

$$dr(\mathbf{x},t)\exp^{\phi(\mathbf{x},\mathbf{y},t)} = A^* r(\mathbf{x},t)\exp^{\phi(\mathbf{x},\mathbf{y},t)} dt + (d\phi(\mathbf{x},\mathbf{y},t)) \, r(\mathbf{x},t) \exp^{\phi(\mathbf{x},\mathbf{y},t)}$$

$$\exp^{\phi(\mathbf{x},\mathbf{y},t)} = dr(\mathbf{x},t) + r(\mathbf{x},t) d \exp^{\phi(\mathbf{x},\mathbf{y},t)} \;=\; A^* r(\mathbf{x},t) \exp^{\phi(\mathbf{x},\mathbf{y},t)} dt$$
$$+ (d\phi(\mathbf{x},\mathbf{y},t)) \, r(\mathbf{x},t) \exp^{\phi(\mathbf{x},\mathbf{y},t)}$$

$$dr(\mathbf{x},t) = \exp^{-\phi(\mathbf{x},\mathbf{y},t)} [\; A^*r(\mathbf{x},t)\exp^{\phi(\mathbf{x},\mathbf{y},t)} dt + (d\phi(\mathbf{x},\mathbf{y},t))\, r(\mathbf{x},t)\exp^{\phi(\mathbf{x},\mathbf{y},t)}$$
$$-r(\mathbf{x},t)\left(d\exp^{\phi(\mathbf{x},\mathbf{y},t)}\right)]$$

$$dr(\mathbf{x},t) = \exp^{-\phi(\mathbf{x},\mathbf{y},t)} [\; A^*r(\mathbf{x},t)\exp^{\phi(\mathbf{x},\mathbf{y},t)} dt + (d\phi(\mathbf{x},\mathbf{y},t))\, r(\mathbf{x},t)\exp^{\phi(\mathbf{x},\mathbf{y},t)}$$
$$-(d[\phi(\mathbf{x},\mathbf{y},t)])\, r(\mathbf{x},t)\exp^{\phi(\mathbf{x},\mathbf{y},t)}]$$

$$dr(\mathbf{x},t) = \exp^{-\phi(\mathbf{x},\mathbf{y},t)} \left[A^*r(\mathbf{x},t)\exp^{\phi(\mathbf{x},\mathbf{y},t)} \right] dt$$

or

$$dr(\mathbf{x},t) = \exp^{-\phi(\mathbf{x},\mathbf{y},t)} A^* \exp^{\phi(\mathbf{x},\mathbf{y},t)} r(\mathbf{x},t) dt \;\; .$$

This is a classic parabolic partial differential equation. Using results from semi-group theory and differential equations the solution is given by

$$r(\mathbf{x},t) = \exp^{-\phi(\mathbf{x},\mathbf{y},t)} \exp^{A^*t} \exp^{\phi(\mathbf{x},\mathbf{y},t)} u(\mathbf{x},0)$$

so

$$
\begin{aligned}
u(\mathbf{x},t) &= \exp^{\phi(\mathbf{x},\mathbf{y},t)} r(\mathbf{x},t) = \exp^{A^*t} \exp^{\phi(\mathbf{x},\mathbf{y},t)} u(\mathbf{x},0)\\
&= \exp^{A^*t} \exp^{h^T dy - \frac{1}{2}\|h\|^2 t + C(\mathbf{x})t} u(\mathbf{x},0) \;\; .
\end{aligned}
\tag{2.3}
$$

2.2 Calculation of Likelihood Ratio Λ

From the introduction we have for the single hypothesis versus noise case

$$\Lambda_t(\mathbf{y}) = \exp\left\{ \int_0^t \hat{h}_s^T dy_s - \frac{1}{2}\int_0^t \|\hat{h}_s\|^2 ds \right\}$$

and for the general hypothesis case

$$\Lambda_t(\mathbf{y}) = \exp\left\{ \int_0^t \left(\hat{h}_s^1 - \hat{h}_s^0\right)^T dy_s - \frac{1}{2}\int_0^t \left(\|\hat{h}_s^1\|^2 - \|\hat{h}_s^0\|^2\right) ds \right\} \;\; .$$

To calculate Λ_t from the above formulas it is necessary to find $\hat{h}_t^i$, the conditional expectation of h_t^i. We will solve the Zakai equation for $u_i(\mathbf{x},t)$ the unnormalized conditional density assuming hypothesis i, then taking the expectation of h_t^i with

7

respect to the normalized conditional density we can calculate Λ_t. However, we will show that

$$\Lambda_t(\mathbf{y}) = \int_{I\!\!R^n} u_1(\mathbf{x}, t) dx$$

in the single hypothesis case and

$$\Lambda_t(\mathbf{y}) = \frac{\int_{I\!\!R^n} u_1(\mathbf{x}, t) dx}{\int_{I\!\!R^n} u_0(\mathbf{x}, t) dx}$$

in the general hypothesis case. By using the above formulas to calculate Λ_t we avoid the additional approximation in calculating $\hat{h}_t^i$.

Lemma 2

$$\Lambda_t(\mathbf{y}) = \exp\left\{\int_0^t \hat{h}_s^T dy_s - \frac{1}{2}\int_0^t \|\hat{h}_s\|^2 ds\right\}$$

(for the single hypothesis versus noise case) satisfies the stochastic differential equation

$$d\Lambda_t(\mathbf{y}) = \Lambda_t(\mathbf{y})\hat{h}^T dy \quad.$$

proof: *Changing notation let $\Lambda_t(\mathbf{y})$ be denoted by $\Lambda(\mathbf{y}, t)$. We use ' to denote differentiation and the corresponding subscripts will denote the variable with respect to which the differentiation is performed. Let*

$$z(t) = \int_0^t \hat{h}_s^T dy_s - \frac{1}{2}\int_0^t \|\hat{h}_s\|^2 ds$$

then

$$dz(t) = \hat{h}^T dy - \frac{1}{2}\|\hat{h}\|^2 dt \quad.$$

Using the fact

$$\begin{aligned} d\mathbf{y} &= h(\mathbf{x})dt + d\mathbf{V} \\ &= \hat{h}(\mathbf{x})dt + d\tilde{\mathbf{V}} \end{aligned}$$

where $d\tilde{\mathbf{V}}$ is a Brownian motion vector independant of $d\mathbf{V}$ but with the same distribution and is obtained by taking expectations, we get

$$dz(t) = \frac{1}{2}\|\hat{h}\|^2 dt + \hat{h}^T d\tilde{\mathbf{V}} \quad.$$

So $z(t)$ satisfies a stochastic differentail equation. Applying Ito's formula to

$$\Lambda(z(t),t) = \exp\{z(t)\}$$

we get

$$
\begin{aligned}
d\Lambda(z(t),t) &= \left[\Lambda'_t(z(t),t) + \frac{1}{2}||\hat{h}||^2\Lambda'_z(z(t),t) + \frac{1}{2}\hat{h}^T\hat{h}_t\Lambda''_{zz}(z(t),t)\right]dt \\
&\quad + \hat{h}^T\Lambda'_z(z(t),t)d\mathbf{W} \\
&= \left[\frac{1}{2}||\hat{h}||^2\Lambda(z(t),t) + \frac{1}{2}||\hat{h}||^2\Lambda(z(t),t)\right]dt + \hat{h}^T\Lambda(z(t),t)d\mathbf{W} \\
&= ||\hat{h}||^2\Lambda(z(t),t)dt + \hat{h}^T\Lambda(z(t),t)d\mathbf{W} \\
&= \Lambda(z(t),t)\left[||\hat{h}||^2dt + \hat{h}^Td\mathbf{W}\right] \\
&= \Lambda(z(t),t)\left[\hat{h}_t^T\hat{h}dt + \hat{h}^Td\mathbf{W}\right] \\
&= \Lambda(z(t),t)\hat{h}^T\left[\hat{h}dt + d\mathbf{W}\right] \quad .
\end{aligned}
$$

Substituting dy back in for $\hat{h}dt + d\mathbf{W}$ we get

$$d\Lambda(\mathbf{y},t) = \Lambda(\mathbf{y},t)\hat{h}^Tdy \quad .$$

Consider the inner product defined by

$$<g(\mathbf{x}),h(\mathbf{x})> = \int_{I\!\!R^n} g(\mathbf{x})h(\mathbf{x})d\mathbf{x} \quad .$$

For the Zakai equation we have

$$L^*u(\mathbf{x},t) = \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j}[\sigma_{i,j}(\mathbf{x})u(\mathbf{x},t)] - \sum_{i=1}^n \frac{\partial}{\partial x_i}[f_i(\mathbf{x})u(\mathbf{x},t)]$$

with adjoint operator L given by

$$Lu(\mathbf{x},t) = \sum_{i,j=1}^n \sigma_{i,j}(\mathbf{x})\frac{\partial^2}{\partial x_i \partial x_j}[u(\mathbf{x},t)] - \sum_{i=1}^n f_i(\mathbf{x})\frac{\partial}{\partial x_i}[u(\mathbf{x},t)] \quad .$$

So $<L^*g(\mathbf{x}),h(\mathbf{x})> = <g(\mathbf{x}),Lh(\mathbf{x})>$.

From (2.1) we have

$$d<u(\mathbf{x},t),1> = <L^*u(\mathbf{x},t),1>dt + <u(\mathbf{x},t)h^T,1>dy \quad .$$

Then

$$
\begin{aligned}
d < u(\mathbf{x},t), 1 > \; &= \; < u(\mathbf{x},t), L1 > dt + < u(\mathbf{x},t)h^T, 1 > dy \\
&= \; \frac{< u(\mathbf{x},t), h^T >}{< u(\mathbf{x},t), 1 >} < u(\mathbf{x},t), 1 > d\mathbf{y} \\
&= \; < u(\mathbf{x},t), 1 > \hat{h}^T dy
\end{aligned}
$$

with $< u(\mathbf{x},0), 1 >= 1$. So Λ_t and $< u(\mathbf{x},t), 1 >$ satisfy the same stochastic differential equation and are equal P-a.s. So

$$
\begin{aligned}
\Lambda_t \; &= \; < u(\mathbf{x},t), 1 > \\
&= \; \int_{I\!\!R^n} u(\mathbf{x},t) d\mathbf{x} \quad .
\end{aligned}
$$

For the genaral hypothesis case we get

$$
\begin{aligned}
\Lambda_t \; &= \; \frac{< u_1(\mathbf{x},t), 1 >}{< u_0(\mathbf{x},t), 1 >} \\
&= \; \frac{\int_{I\!\!R^n} u_1(\mathbf{x},t) d\mathbf{x}}{\int_{I\!\!R^n} u_0(\mathbf{x},t) d\mathbf{x}} \quad .
\end{aligned}
$$

2.3 Numerical approximation of the Zakai equation

Discretizing (2.3) in time as

$$
U(\mathbf{x},t) = \prod_{l=0}^{\lfloor \frac{t}{\Delta t} \rfloor} \exp^{A^* \Delta t} \exp^{h^T \Delta y_{l\Delta t} - \frac{1}{2}\|h\|^2 \Delta t + C(\mathbf{x})\Delta t} u(\mathbf{x},0) \tag{2.4}
$$

LaVigna shows $lim_{\Delta t \to 0} U(\mathbf{x},t) \to u(\mathbf{x},t)$. Similarly, he shows the approximation

$$
U(\mathbf{x},t) = \prod_{l=0}^{\lfloor \frac{t}{\Delta t} \rfloor} \exp^{h^T \Delta y_{l\Delta t} - \frac{1}{2}\|h\|^2 \Delta t + C(\mathbf{x})\Delta t} \exp^{A^* \Delta t} u(\mathbf{x},0) \tag{2.5}
$$

also converges to $u(\mathbf{x},t)$ as $\Delta t \to 0$.

We will discretize (2.4) and (2.5) in space (ie. with respect to the underlying state $\mathbf{x}$). We will replace the differential operator A^* by matrix operations.

We will use boldface capital letters to denote matricies. We discretize the density $u(\mathbf{x},t)$ as a matrix

$$\begin{bmatrix} u(\mathbf{x}_0,t) & u(\mathbf{x}_0+\Delta\mathbf{x}_2,t) & \cdots & u(\mathbf{x}_0+k\Delta\mathbf{x}_2,t) \\ u(\mathbf{x}_0+\Delta\mathbf{x}_1,t) & u(\mathbf{x}_0+\Delta\mathbf{x}_1+\Delta\mathbf{x}_2,t) & \cdots & u(\mathbf{x}_0+\Delta\mathbf{x}_1+k\Delta\mathbf{x}_2,t) \\ \vdots & \vdots & \ddots & \vdots \\ u(\mathbf{x}_0+j\Delta\mathbf{x}_1,t) & u(\mathbf{x}_0+j\Delta\mathbf{x}_1+\Delta\mathbf{x}_2,t) & \cdots & u(\mathbf{x}_0+j\Delta\mathbf{x}_1+k\Delta\mathbf{x}_2,t) \end{bmatrix}$$

which we will denote by $\mathbf{U}(\mathbf{x},t)$ where $\mathbf{x}_0$ is the left endpoint of the discretization

space.

Consider the first and second order approximations

$$\frac{\partial}{\partial x_i}u(\mathbf{x},t) = \frac{u(x+\Delta\mathbf{x}_i,t)-u(\mathbf{x},t)}{\Delta x_i}$$

$$\frac{\partial}{\partial x_i}u(\mathbf{x},t) = \frac{u(\mathbf{x},t)-u(x+\Delta\mathbf{x}_i,t)}{\Delta x_i}$$

$$\frac{\partial^2}{\partial x_i^2}u(\mathbf{x},t) = \frac{u(x+\Delta\mathbf{x}_i,t)-2u(\mathbf{x},t)+u(x-\Delta\mathbf{x}_i,t)}{\Delta x_i^2}$$

where Δx_i is the discretization increment of the ith component of the state vector

and is scalar and $\Delta\mathbf{x}_i$ is an n dimension vector with Δx_i in the ith position and zero elsewhere. So $\frac{\partial^2}{\partial x_i^2}u(\mathbf{x},t)$ and $\frac{\partial}{\partial x_i}u(\mathbf{x},t)$ are the second and first derivatives of the density with respect to the ith component of the state. Define

$$D_i(\mathbf{x}) = \frac{-2\Delta t}{\Delta x_i^2}\sigma_{i,i}(\mathbf{x}) - \frac{\Delta t}{\Delta x_i}\left| -f_i(\mathbf{x})+2\frac{\partial}{\partial x_i}\sigma_{i,i}(\mathbf{x})\right|$$

$$UD_i(\mathbf{x}) = \frac{\Delta t}{\Delta x_i^2}\sigma_{i,i}(\mathbf{x}) + \frac{\Delta t}{\Delta x_i}\max\left[0,-f_i(\mathbf{x})+2\frac{\partial}{\partial x_i}\sigma_{i,i}(\mathbf{x})\right]$$

$$LD_i(\mathbf{x}) = \frac{\Delta t}{\Delta x_i^2}\sigma_{i,i}(\mathbf{x}) - \frac{\Delta t}{\Delta x_i}\min\left[0,-f_i(\mathbf{x})+2\frac{\partial}{\partial x_i}\sigma_{i,i}(\mathbf{x})\right]$$

where the functions $\max[\cdot,\cdot]$ and $\min[\cdot,\cdot]$ take the maximum or minimum of

their respective arguments. We define the tridiagonal matrices $\mathbf{A}_1^*$ and $\mathbf{A}_2^*$ respectively by

$$\begin{bmatrix} D_1(\mathbf{X}_0) & -D_1(\mathbf{X}_0) & & & \\ LD_1(\mathbf{X}_0+\Delta\mathbf{X}_1) & D_1(\mathbf{X}_0+\Delta\mathbf{X}_1) & UD_1(\mathbf{X}_0+\Delta\mathbf{X}_1) & & \\ & \ddots & \ddots & \ddots & \\ & & LD_1(\mathbf{X}_0+(j-1)\Delta\mathbf{X}_1) & D_1(\mathbf{X}_0+(j-1)\Delta\mathbf{X}_1) & UD_1(\mathbf{X}_0+(j-1)\Delta\mathbf{X}_1) \\ & & & -D_1(\mathbf{X}_0+j\Delta\mathbf{X}_1) & D_1(\mathbf{X}_0+j\Delta\mathbf{X}_1) \end{bmatrix}$$

$$
\begin{bmatrix}
D_2(x_0,t) & -D_2(x_0) & & \\
LD_2(x_0+\Delta x_2) & D_2(x_0+\Delta x_2) & UD_2(x_0+\Delta x_2) & \\
\ddots & \ddots & \ddots & \\
& LD_2(x_0+(j-1)\Delta x_2) & D_2(x_0+(j-1)\Delta x_2) & UD_2(x_0+(j-1)\Delta x_2) \\
& & -D_2(x_0+j\Delta x_2) & D_2(x_0+j\Delta x_2)
\end{bmatrix} .
$$

With the further assumption that $f_i(\mathbf{x})$ only depends on the ith component of the state vector we can rewrite (2.4) and (2.5) as

$$
\mathbf{U}(\mathbf{x},t) = \prod_{l=0}^{\lfloor \frac{t}{\Delta t} \rfloor} \exp^{A_1 \Delta t} \left[\mathbf{E}(l\Delta t) \otimes \mathbf{U}(\mathbf{x},l\Delta t) \right] \exp^{A_2 \Delta t} \tag{2.6}
$$

and

$$
\mathbf{U}(\mathbf{x},t) = \prod_{l=0}^{\lfloor \frac{t}{\Delta t} \rfloor} \mathbf{E}(l\Delta t) \otimes \exp^{A_1 \Delta t} \mathbf{U}(\mathbf{x},l\Delta t) \exp^{A_2 \Delta t} \tag{2.7}
$$

respectively. The symbol $\otimes$ is pointwise matrix multiplication and $\mathbf{I}$ is either an $j \times j$ or a $k \times k$ identity matrix. $\mathbf{E}(t)$ is defined as

$$
\exp \left[h^T \Delta \mathbf{y}_t \right]
\begin{bmatrix}
\alpha(x_0) & \alpha(x_0+\Delta x_2) & \cdots & \alpha(x_0+k\Delta x_2) \\
\alpha(x_0+\Delta x_1) & \alpha(x_0+\Delta x_1+\Delta x_2) & \cdots & \alpha(x_0+\Delta x_1+k\Delta x_2) \\
\vdots & \vdots & \ddots & \vdots \\
\alpha(x_0+j\Delta x_1) & \alpha(x_0+j\Delta x_1+\Delta x_2) & \cdots & \alpha(x_0+j\Delta x_1+k\Delta x_2)
\end{bmatrix}
$$

where $\alpha(\mathbf{x},t) = \exp \left[-\frac{1}{2}\|h\|^2 \Delta t + C(\mathbf{x})\Delta t \right]$. Note that in discretizing the equations in space the operation A_i^* for $i = 1$ became matrix multiplication on the left and for $i = 2$ matrix multiplication on the right. If $f_i(\mathbf{x})$ is not assumed to depend only on the ith component of the state vector then (2.4) and (2.5) can still be implemented using matrices, however, $\mathbf{A}_1^*$ changes for each column of the discretized density and $\mathbf{A}_2^*$ changes for each row of the discretized density.

We know

$$
\exp^{A_i \Delta t} = \sum_{n=0}^{\infty} \frac{(A_i \Delta t)^n}{n!} .
$$

We are interested in approximating the exponential of a matrix by a truncated sum. If the elements of $A_i \Delta t \ll 1$, which we can guarantee by choosing Δt small enough, then

$$
\exp^{A_i \Delta t} \approx \mathbf{I} + A_i \Delta t .
$$

12

By using only first and second order approximations for the first and second order derivatives of the density with respect to the state we have the row of A_i summing to zero. Additionally, choosing Δt such that the elements of $A_i \Delta t \ll 1$ we have ensured the matrix $I + A_i \Delta t$ is well conditioned and invertible. We can implement several different schemes for (2.6) and (2.7). We can rewrite one step of (2.6) as

$$U(x, t + \Delta t) = (I + \Delta t A_1^*) \left[E(t) \otimes U(x, t) \right] (I + \Delta t A_2^*) \qquad (2.8)$$

which we will denote as the explicit discretization. We can implement a two step approximation of (2.6) by

$$(I - \Delta t A_1^*) U(x, t + \Delta t) = \left[E(t) \otimes U(x, t) \right] (I + \Delta t A_2^*) \qquad (2.9)$$

$$U(x, t + 2\Delta t) (I - \Delta t A_2^*) = (I + \Delta t A_1^*) \left[E(t) \otimes U(x, t) \right] \qquad (2.10)$$

which we will denote as the mixed scheme. Finally, we can implement the approximation

$$(I - \Delta t A_1^*) U(x, t + \Delta t) (I - \Delta t A_2^*) = \left[E(t) \otimes U(x, t) \right] \qquad (2.11)$$

which we denote as the implicit scheme. (2.8) is called explicit because the density $U(x, t + \Delta t)$ is explicitly defined interms of the density $U(x, t)$. (2.11) is called implicit because each point in the density $U(x, t + \Delta t)$ is defined implicitly by one or more points in the density $U(x, t + \Delta t)$ as well as $U(x, t)$ and a set of simultaneous equations must be solved to obtain $U(x, t + \Delta t)$. The scheme corresponding to (2.9) and (2.10) having components of both the explicit and implicit schemes is denoted as the mixed scheme. For implementing (2.7) we again get similar result with the pointwise multiplication by $E(t)$ done after matrix multiplication of $I + \Delta t A_i^*$ and $U(x, t)$.

Several approximations were used, in arriving at a discretization scheme for (2.4) or (2.5), which warrant further discussion. We approximated the exponential of a matrix by using only two terms of the Taylor expansion. In order

for the approximation to be accurate it is necessary that the elements of the matrix be small. The only term which is free to be set is Δt but making Δt small requires more computations to process a fixed time block of data. Alternately, a better approximation for the exponential can be used by including more terms of the expansion but for each additional term the matrices A_i^* have an additional nonzero upper and lower diagonal which again increases computational complexity and A_i^* can nolonger be guaranteed to be a well conditioned matrix. Another approximation used was the first and second order approximations for the first and second order derivatives of the density $u(\mathbf{x}, t)$. Higher order approximations can be used but result in additional nonzero upper and lower diagonals in A_i^* which again increases computational complexity and the matrix $I + A_i \Delta t$ can not be guaranteed to be well conditioned. There is also a question of which scheme to implement to approximate the conditional density $u(\mathbf{x}, t)$. The implicit schemes require more computations than the explicit schemes but the implicit schemes exhibit better numerical properties as Δt becomes large. However, since we require Δt to be small the implicit, explicit and mixed schemes are essentially equivalent with respect to numerical stability. There is also the approximation of the conditional density $u(\mathbf{x}, t)$by the matrix $U(\mathbf{x}, t)$. If the discretization is very coarse $U(\mathbf{x}, t)$ will be a poor approximation of the density $u(\mathbf{x}, t)$but the more fine it becomes the more calculations per step in time are needed. Finally, there is a question as to which aproximation, (2.6) or (2.7), is better to implement. Some of these considerations have been studied for particular cases of interest to the Navy. These cases impose additional constraints on the implementation scheme used and are discussed in Chapter 3 and Chapter 4.

2.4   Implementation

The discretized schemes were implemented using MACSYMA and FORTRAN. MACSYMA is a symbolic math manipulation package written by the MIT Laboratory for Computer Science. It is able to generate legal FORTRAN code which was used to implement the actual numerical calculations. The use of MACSYMA gave significant flexibility in entering models by allowing symbolic calculations including differentiation required to implement the numerical solution. It also allowed for easy changes to the discretization of the state space. After specifying the models and state space discretization, FORTRAN code is automatically generated which simulates the diffusion (using first order difference equations) and solves the Zakai equation. The code makes use of routines from LINPACK, a set of FORTRAN subroutines for doing linear algebra. The FORTRAN allows entry of any constant parameters at the time of execution.

Presently the MACSYMA code is running on a Texas Instruments Explorer and the FORTRAN code is then transferred to a VMS or Unix machine and executed.

2.5   Comparison of discretization schemes

We presented five possible discretization schemes for the solution to the Zakai equation. For (2.4) there are the implicit, mixed, and explicit schemes. For (2.5) there are the mixed, and explicit schemes. Note that the implicit scheme for (2.5) would be the same as for (2.4). This section gives results comparing the numerical approximation of the Zakai equation by (2.6) or (2.7) and the use of the implicit, mixed of explicit scheme. A comparison for one particular case is shown in figures 2.1 to 2.8. These figures correspond to a lognormal model versus a Rayleigh model with the Rayleigh decorrelation time one tenth that of the lognormal, (reference tables 4.1 and 4.2). The use of these particular models are discussed in Chapter 3. In figures 2.1 to 2.4 we show the log

likelihood ratios given hypothesis 0 data and in figures 2.5 to 2.8 we show the log likelihood ratio given hypothesis 1 data for one particular set of parameters. We see that in all cases the log likelihood ratios are very similar. Again, this is partly due to the small time discretization steps which was necessary for the approximation to the exponential of a matrix to be accurate. All the plots of the log likelihood ratio are compared with the implicit discretization of (2.4) which is given by (2.11). From Monte Carlo simulations presented in Chapter 5, using the implicit discretization and the mixed discretization of (2.4), we conclude that the results are nearly identical in the number of correct decisions and the decision times which supports the conclusion suggested by figures 2.1 to 2.8 that for the given models the approximation scheme behave nearly identically.
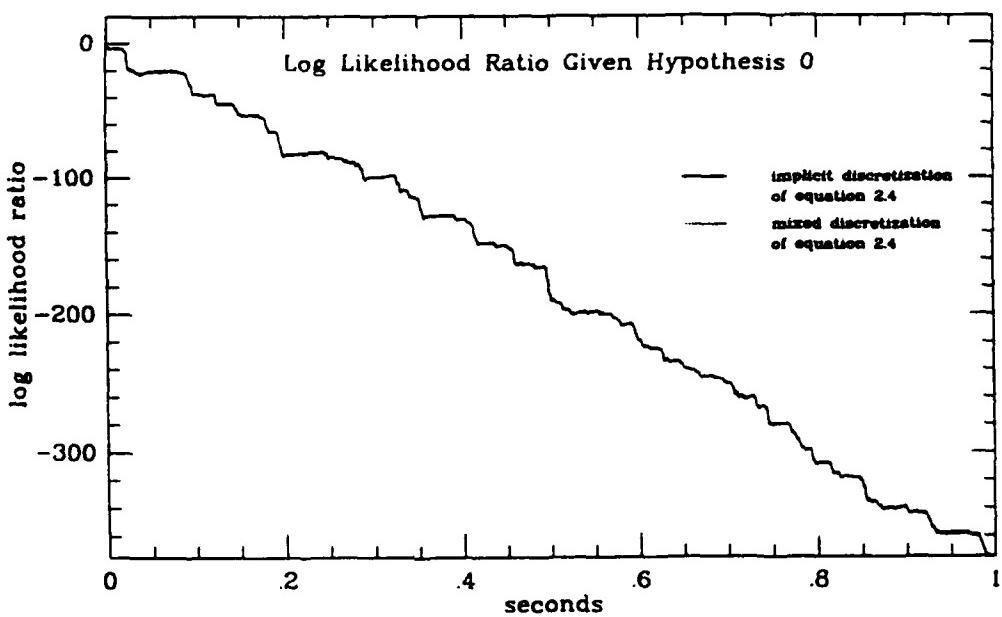
Figure 2.1: Mixed discretization of 2.4 under hypothesis 0, SNR=5
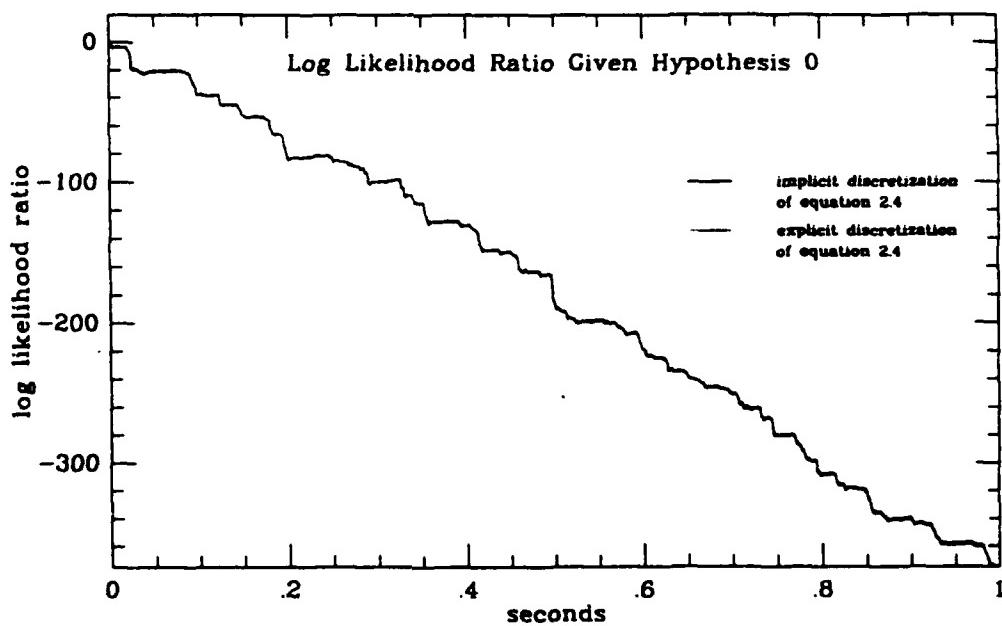

Figure 2.2: Explicit discretization of 2.4 under hypothesis 0, SNR=5

17
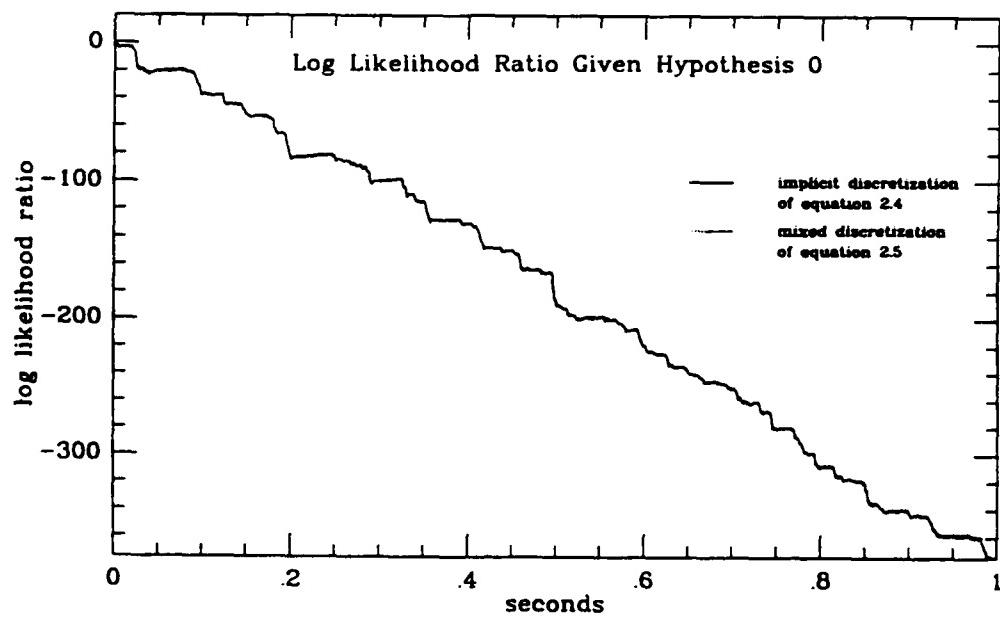
Figure 2.3: Mixed discretization of 2.5 under hypothesis 0, SNR=5
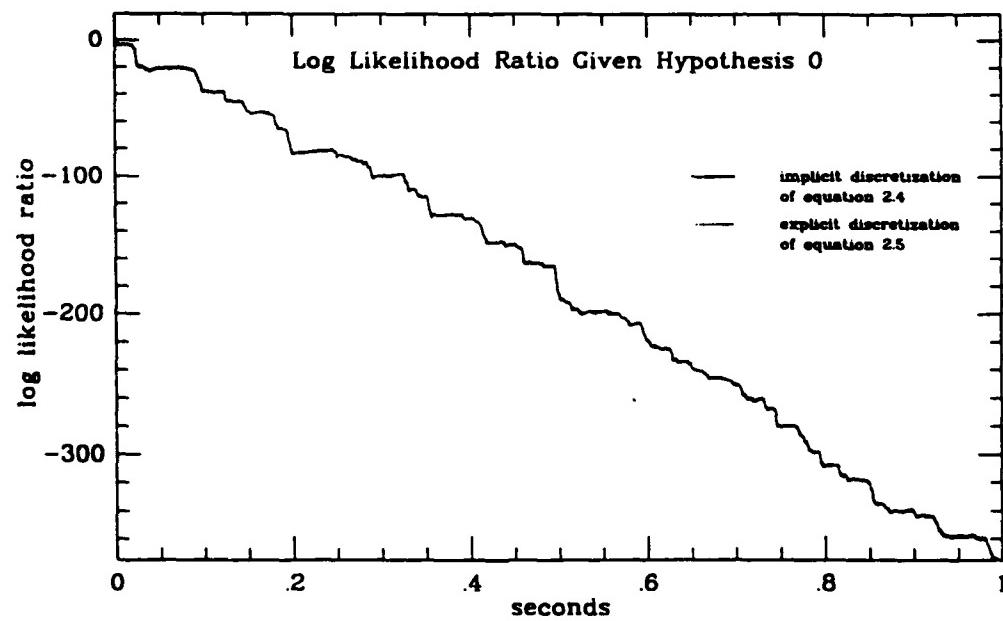
Figure 2.4: Explicit discretization of 2.5 under hypothesis 0, SNR=5
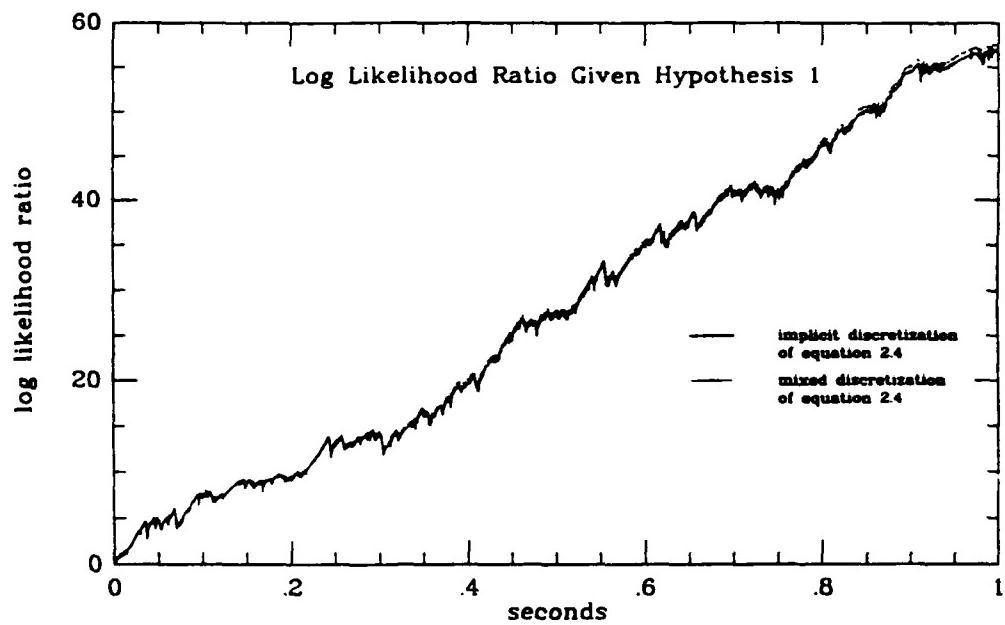
Figure 2.5: Mixed discretization of 2.4 under hypothesis 1, SNR=5
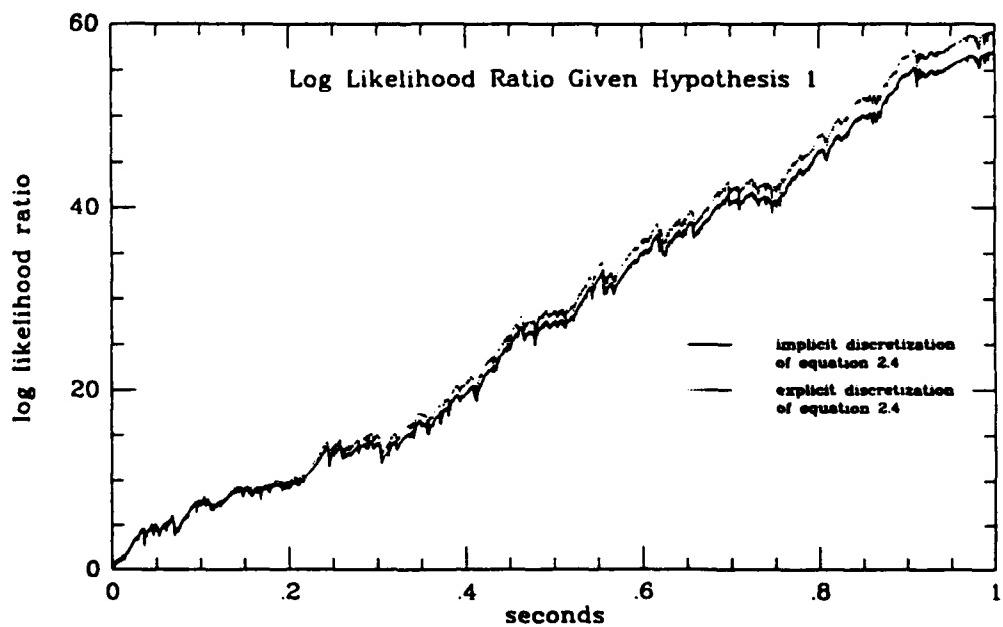

Figure 2.6: Explicit discretization of 2.4 under hypothesis 1, SNR=5
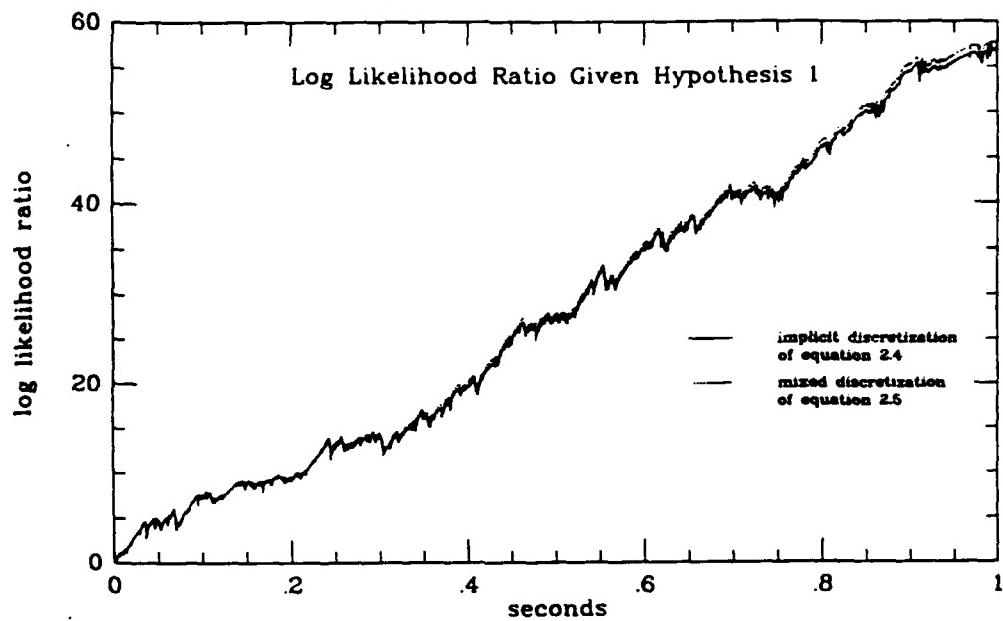
Figure 2.7: Mixed discretization of 2.5 under hypothesis 1, SNR=5
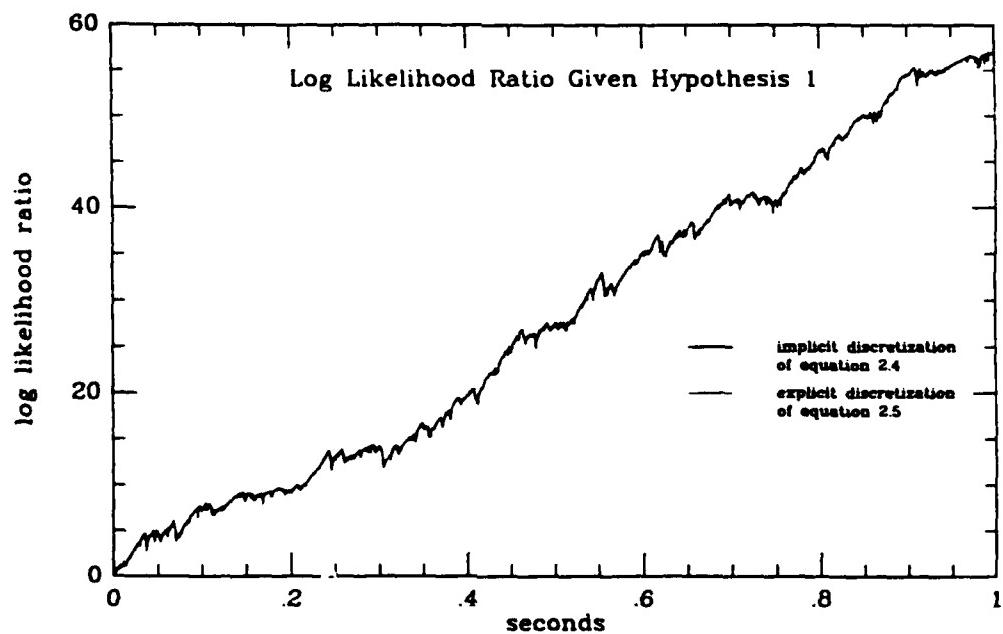

Figure 2.8: Explicit discretization of 2.5 under hypothesis 1, SNR=5

distribution.

The typical radar system is composed of an IF amplifier followed by a second detector and a video amplifier. The second detector and the video amplifier form an envelope detector. If the inphase and quadrature components of the signal entering the second detector are independent and have a Gaussian distribution then Rice has shown the density of the envelope to be Rayleigh (amplitude) and the power will be distributed exponentially (Rayleigh Power).

The Rayleigh distribution has also been shown to model sea clutter when the resolution cell (ie. the area illuminated by the radar pulse) is large relative to the water wavelength. It has also been seen that sea clutter with high resolution radar will often have heavier tails which can better be represented by a lognormal density function.

In addition to the amplitude distribution of the chaff and clutter returns, the power spectrum is also very important. It gives information as to the rate at which the radar cross section changes. Early models of the spectrum were assumed to be Gaussian which gives an autocovariance that is also Gaussian. More recent measurements performed in the 1960's and 1970's showed power spectrums of the form

$$\frac{A}{1 + \left(\frac{f}{f_c}\right)^n}$$

to be more appropriate, where

  A is the mean value of the power density

  f_c is the clutter spectrum half-power frequency

  f is frequency

  n is a positive real number.

According to Barton, the actual spectrum can generally be fitted with a "band-limited white noise spectrum, extending from zero to infinity, but with steadily decreasing amplitude above (a) given frequency" [Bar64,p.82], which we denote by f_c (the half-power frequency). Then the actual spectrum can be well approx-

imated by

$$\frac{A}{1 + \left(\frac{f}{f_c}\right)^2}$$

The autocovariance function is given by

$$R(t) = A\pi f_c \exp\left\{-2\pi f_c|t|\right\}$$

which results in exponential decorrelation with time constant $\frac{1}{2\pi f_c}$.

In the case of a ship the scattering elements are much more complex and most of the modelling has been based on empirical considerations. "The lognormal distribution has been used to model scattering from highly directive reflectors when viewed from random aspects" [Bla86,p.77], such as randomly oriented flat plates. corner reflectors, and antennas. "It has become popular to use the lognormal model to describe large metal objects of irregular shape, especially ocean vessels with complex superstructures" [MM73,p.67]. We do not wish to imply, however, that the lognormal distribution can model all RCS returns from ships. In addition to the lognormal distribution, radar cross section (RCS) returns have also been observed to be distributed Rayleigh power, Rayleigh amplitude, chi-square with two and four degrees of freedom, and Rician. Lognormal statistics tend to appear at major aspects of ships with dominant scatterers, especially on the larger ships such as carriers and battleships.

As in the case of chaff and clutter, it has been observed "that the power spectral density of the ship RCS fluctuation due to deterministic and random azimuth(aspect), pitch, and roll motion can be very well approximated by a power spectral density of the form" [Bar80,p.30]

$$\frac{A}{1 + \left(\frac{f}{f_b}\right)^2}$$

which yields the autocovariance function

$$R(t) = A\pi f_b \exp\left\{-2\pi f_b|t|\right\}$$

which is exponential decorrelation with time constant $\frac{1}{2\pi f_b}$. It has been observed that the decorrelation time for a ship tends to be larger than for chaff or clutter

Chapter 3

Model Definition and Parameter Estimation

3.1 Specification of model

We are interested in the performance of the detector for the previously described discretization schemes and as such we need specific models. The Navy is interested in the application of the resulting detector to radar problems. Based on these radar problems we choose specific signal and noise models. We will look at the ship versus decoy problem.

One common decoy used is chaff. Chaff is the code name used during WW II to refer to metallic dipoles (strips of light metal foil) dropped or launched into the air to confuse enemy radar by presenting a large radar return. If these dipoles are half the length of the radar wavelength they will resonate and give large radar returns. Since the chaff cloud consists of a large number of scatterers of which no single one dominates, the phase and amplitude variations of the individual scattering elements can be considered independent. By application of the Central Limit theorem to the phasor summation of the RF voltages induced in the receiving antenna by the reflection of the individual scatterers it leads to the conclusion that the inphase and quadrature components have a Gaussian

which is supported by the fact that the individual scattering elements of a ship are fixed relative to each other and the azimuth, pitch, and roll motion is slow relative the orientation changes of the dipoles in the falling chaff or the sea surface.

The Rayleigh and lognormal radar returns can be modeled using diffusions which are given by

$$dx_t = f^i(x_t)dt + g^i(x_t)dW_t^i$$
$$dy_t = h^i(x_t)dt + \gamma dV_t \ .$$

We will drop the time subscript and use subscripts to denote elements in vectors or matrices. Under hypothesis 0 we will assume that

$$f^0(x) = a \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$g^0(x) = \begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix}$$
$$h^0(x) = [\, c(x_1^2 + x_2^2)^{\frac{1}{2}} \,],$$

then $h^0(x)$ has Rayleigh amplitude statistics. As for hypothesis 1, we assume that

$$f^1(x) = [\, qx_1 \,]$$

$$g^1(x) = [\, r \,]$$

$$h^1(x) = [\, s \exp\{x_1\} \,],$$

then $h^1(x)$ is distributed according to a lognormal density. The parameters a, b, c, q , r, s, and γ are constants. The term γdV_t models thermal noise in the receiver. The constant γ must be nonzero or else the likelihood ratio Λ_t becomes undefined which results in numerical problems in the discretization.

We will wish to test the detector with known parameter values and with estimated values. The parameters we need to estimate are the constants a, b, c, γ in the Rayleigh case and q, r, s, γ in the lognormal case. In the case $\gamma \neq 1$,

as explained earlier, the term γ becomes a renormalizing term. When a and b (q and r) are independent of time the underlying state is a correlated Gaussian path. For the discretized version of the differential equation the parameters a and b (q and r) can then be related to the parameters ρ_G the correlation and σ_G stationary variance of the underlying Gaussian (appendix D) by

$$a = \frac{\rho_G - 1}{\Delta t}$$

$$b = \sigma_G \left(\frac{1 - (1 + a\Delta t)^2}{\Delta t} \right)^{\frac{1}{2}}$$

where $\rho_G = \exp\left(\frac{-\Delta t}{t_G}\right)$ and t_G is the decorrelation time constant for the underlying Gaussian.

3.2   Parameter estimation: lognormal

One very common estimator is the maximum likelihood estimator (MLE). To calculate the MLE of the parameters for the discretized version of the solution to the Zakai equation we need to find the distribution of the output data. To generate the lognormal or Rayleigh distribution given the form of the models we first look at the distribution for the underlying Gaussian. With the requirement that $\frac{-2}{\Delta t} < a < 0$ (or q respectively), where Δt is the time discretization used, then the underlying Gaussian has a stationary distribution with zero mean, variance denoted σ_G^2, and exponential decorrelation ie. $R(k) = \exp^{\frac{-k\Delta t}{t_G}} = \rho_G^k$. We can write the distribution for the underlying Gaussian in closed form for both the lognormal and the Rayleigh cases, refer to appendix A.

For the lognormal case $h^1(\mathbf{x})$ is an invertible transformation, so we can use the Jacobian change of variables formula to determine the distribution for $h^1(\mathbf{x})$. Denote $y_i = h^1(\mathbf{x}_i)$ then we can write the distibution for $\mathbf{y} = [y_1, y_2, ..., y_n]$ as

$$f(\mathbf{y}) \ = \ \frac{1}{(2\pi)^{(n/2)} \sigma_G^n (1 - \rho_G^2)^{\frac{(n-1)}{2}} \prod_{i=1}^{n} y_i}$$

$$\exp\left[\frac{-1}{2\sigma_G^2(1-\rho_G^2)}\left[\sum_{i=1}^{n}[i]^2 - 2\rho_G\sum_{i=1}^{n-1}[i][i+1] + \rho_G^2\sum_{i=2}^{n-1}[i]^2\right]\right]$$

where

$$[i] = \ln\left(\frac{y_i}{s}\right) - \mu = \ln(y_i) - (\ln(s) + \mu) \quad .$$

To calculate the distribution of the observation ($h^i(\mathbf{x})$ plus additive noise) in closed form is not possible so one cannot find the MLE, however, if the signal to noise ration (SNR) is very large the ML estimator derived assuming no noise may be a good estimator. The ML estimators ,assuming no noise,for σ_G^2 and $\ln(s) - \mu$ with respect to ρ_G are given by

$$\ln(s) - \mu = \frac{\ln(y_1) + \ln(y_n) + (1-\rho_G)\sum_{i=2}^{n-1}\ln(y_i)}{2 + (n-2)(1-\rho_G)}$$

$$\sigma_G^2 = \frac{1}{n(1-\rho_G^2)}\left[\sum_{i=1}^{n}[i]^2 - 2\rho_G\sum_{i=1}^{n-1}[i][i+1] + \rho_G^2\sum_{i=2}^{n-1}[i]^2\right]$$

where

$$[i] = \ln(y_i) - (\ln(s) + \mu) \quad .$$

The ML estimator for ρ_G cannot be solved for explicitly, however, we are only interested in solutions in the interval $(-1, 1)$ and can easily implement an algorithm to find the roots of the derivative of the density with respect to ρ_G in this interval. Differentiating the lognormal density with respect to ρ_G and setting it equal to zero we get

$$0 = (n-1)\sigma_G^2\rho_G(1-\rho_G^2) - \rho_G\left[T_1 - 2\rho_G T_2 + \rho_G^2 T_3\right] + (1-\rho_G^2)\left[T_2 - \rho T_3\right] \quad (3.1)$$

with

$$T_1 = \sum_{i=1}^{n}\ln^2(y_i) - 2\overline{\mu}\sum_{i=1}^{n}\ln(y_i) + n\overline{\mu}^2$$

$$T_2 = \sum_{i=1}^{n-1}\ln(y_i)\ln(y_{i+1}) - \overline{\mu}\left[\sum_{i=1}^{n-1}\ln(y_i) + \sum_{i=2}^{n}\ln(y_i)\right](n-1)\overline{\mu}^2$$

$$T_3 = \sum_{i=2}^{n-1}\ln^2(y_i) - 2\overline{\mu}\sum_{i=2}^{n-1}\ln(y_i) + (n-2)\overline{\mu}^2$$

which depend on ρ_G thru $\bar{\mu}$, the ML estimate for $\ln(s) + \mu$ for a particular value of ρ_G. By finding the root of (3.1) we obtain an estimate for ρ_G and in turn an estimate for $\ln(s) + \mu$ and σ_G^2. Note that the ML estimator for σ_G^2 can also be written as

$$\sigma_G^2 = \frac{1}{n(1 - \rho_G^2)} \left[T_1 - 2\rho_G T_2 + \rho_G^2 T_3 \right] .$$

These estimates are not very good when any significant amount of additive white noise in the output equation is present, particularly for ρ_G. As will be seen later the detector becomes numerically unstable if the SNR is too large, ie. small noise power relative to signal power, so the ML estimates assuming no noise are not useful in this case. An alternate estimation scheme would be to estimate ρ_G and then use the ML estimator for correlated lognormal data or the ML estimators for independent data which are

$$\ln(s) + \mu = \frac{1}{n} \sum_{i=1}^{n} \ln(y_i)$$

$$\sigma_G^2 = \frac{1}{n} \sum_{i=1}^{n} \ln(y_i) - (\ln(s) + \mu) .$$

We can estimate ρ_G from the covariance function of the output data. If the underlying state is Gaussian with an exponential covariance function then the lognormal sample path has a covariance function which is approximately exponential [Bar80]. Assuming an exponential covariance function for the lognormal signal we match an exponential to $R_o(k)$,$k > 0$(o denoting observation data). From this we get an estimate for the decorrelation time of the lognormal data which we will denote as t_l. Then we use the approximation

$$\frac{1}{t_G} = \frac{1}{t_l} \left[\ln(\sigma_G^2) - \ln \left[\ln \left(1 + \frac{\exp(\sigma_G^2) - 1}{\exp(1)} \right) \right] \right]$$

from [Bar80] to get an estimate for t_G and thus ρ_G. We justify the estimate for t_l by the fact that the lognormal data without additive white noise has a covariance function which is approximately exponential , the white noise covariance function is a Delta function, and since the lognormal data and the white

noise are independent the covariance function of the observation is the sum of the covariance function of the lognormal signal and the covariance function of the noise. Figures 3.1, 3.3 and 3.5 are plots of lognormal model sample paths without noise and with additive noise for SNR of 20, 5 and 2. Figures 3.2, 3.4 and 3.6 are plots of the normalized covariance functions of the signal without noise (solid), the signal plus noise (short and long dash) $R_o(k), k \geq 0$, signal plus noise with the first term removed(short dash) $R_o(k), k > 0$, and an exponential function (long dash). In figures 3.10 to 3.15 we have similar plots for the Rayleigh model for SNR=5 and decorrelation times $t_r = \frac{t_l}{100}, \frac{t_l}{10}, t_l$. We see that the covariance function does exhibit the desired behavior of a δ function at zero for the additive noise plus an approximately exponential function for $R_o(k), k > 0$. An additional observation is that as the SNR becomes smaller (more noise) the exponential behavior of the covariance function, $R_o(k), k > 0$, becomes more corrupted due to the noise covariance function not being a true δ function, especially for the lognormal case where the stationary variance is smaller than for the Rayleigh case and the signal begins to look like a mean which is removed by the covariance function.

Another possible estimator for $\ln(s)+\mu$ and σ_G^2 would be the mean to median ratio of i.i.d. data (reference appendix B). If we assume the underlying state has zero mean then

$$\frac{\text{mean}}{\text{median}} = \exp \frac{\sigma_G^2}{2}$$

$$\text{median} = s \exp \mu = s \quad .$$

Finally, from Huber we have a robust estimator for independent identically distributed Gaussian data known as the Median Absolute Deviation (MAD) estimator [Hub81]. By taking the logarithm of independent noisy data values (rejecting any negative observations) we have corrupted iid Gaussian data. Let $z_i = \ln y_i$ then

$$\ln(s) + \mu = \text{median}\{z_i\}$$

$$\sigma_G^2 = [\text{median}\{|z_i - \text{median}\{z_i\}|\}]^2$$

We again estimate ρ_G from the covariance function.

3.3 Parameter estimation: Rayleigh

For the Rayleigh case if we set $y_i = h^0(\mathbf{x}_i)$ we have the distribution of $\mathbf{y} = [y_1, y_2, ..., y_n]$ as

$$f(\mathbf{y}) = \frac{\prod_{i=1}^n y_i}{\sigma_G^{2n} c^{2n} (1 - \rho_G^2)^{(n-1)}} \exp\left[\sum_{i=1}^n \alpha_i y_i^2\right] \prod_{i=1}^{n-1} I_0 \left[\beta y_i y_{i+1}\right]$$

where

$$\alpha_i = \begin{cases} \frac{-(1+\rho_G^2)}{2\sigma_G^2 c^2 (1-\rho_G^2)} & i = 2, ..., n-1 \\ \frac{-1}{2\sigma_G^2 c^2 (1-\rho_G^2)} & i = 1. n \end{cases}$$

$$\beta = \frac{\rho_G}{\sigma_G^2 c^2 (1 - \rho_G^2)}$$

and I_0 is the modified zero order Bessel function. It is not possible to estimate c and σ_G independently since c and σ_G alway appear together. We can only estimate the product $c\sigma_G$. Since we can only estimate the product of c and σ_G we arbitrarily set the stationary variance of the Rayleigh model to the stationary variance of the lognormal model when we implement the detector.

The ML estimator for $c\sigma_G$ and ρ_G is the solution to the equations

$$0 = 2nc^2\sigma_G^2(1 - \rho_G^2) - \left(y_1^2 + y_n^2 + (1 + \rho^2) \sum_{i=2}^{n-1} y_i^2\right)$$

$$+ \rho_G \sum_{i=1}^{n-1} y_i y_{i+1} \frac{I_1\left(\frac{\rho_G}{c^2\sigma_G^2(1-\rho_G^2)} y_i y_{i+1}\right)}{I_0\left(\frac{\rho_G}{c^2\sigma_G^2(1-\rho_G^2)} y_i y_{i+1}\right)}$$

and

$$0 = 2\rho_G(n-1)(1 - \rho_G^2) - \left(y_1^2 + y_n^2 + 2\sum_{i=2}^{n-1} y_i^2\right)$$

$$+ (1 + \rho_G^2) \sum_{i=1}^{n-1} y_i y_{i+1} \frac{I_1\left(\frac{\rho_G}{c^2\sigma_G^2(1-\rho_G^2)} y_i y_{i+1}\right)}{I_0\left(\frac{\rho_G}{c^2\sigma_G^2(1-\rho_G^2)} y_i y_{i+1}\right)}$$

where I_1 is the modified first order Bessel function. Because of the poor results for the ML estimator for the lognormal case when noise is presen, and the fact
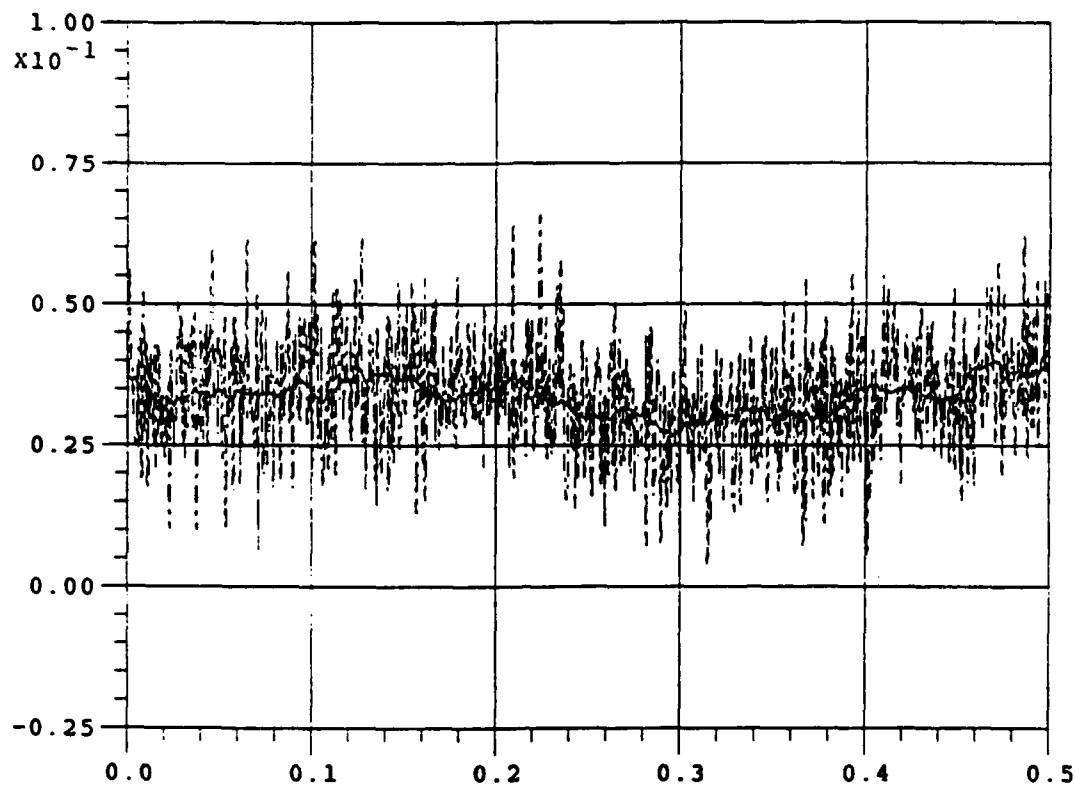
Figure 3.1: Lognormal sample path, SNR=20
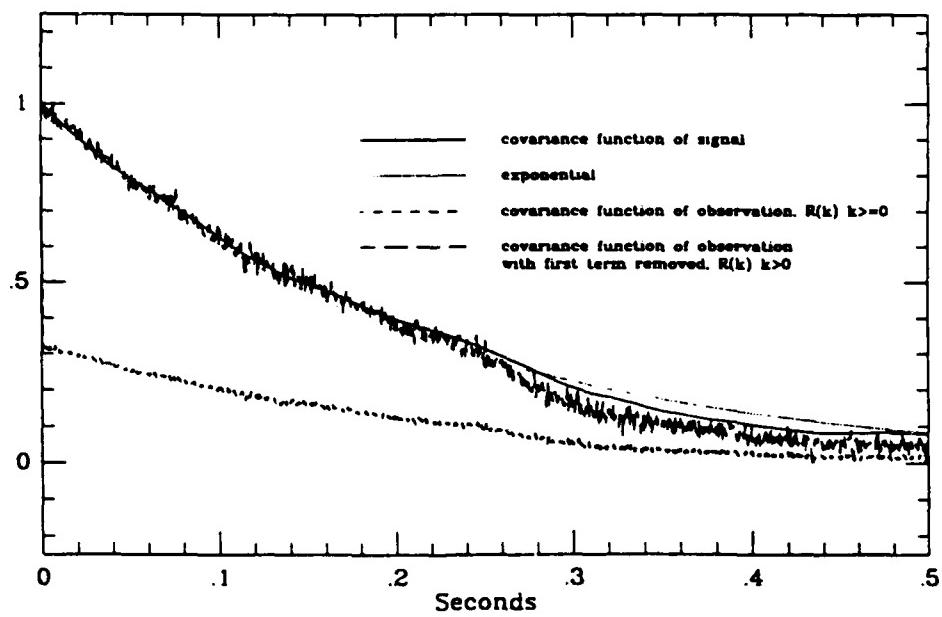
Figure 3.2: Covariance functions for Lognormal sample path, SNR=20
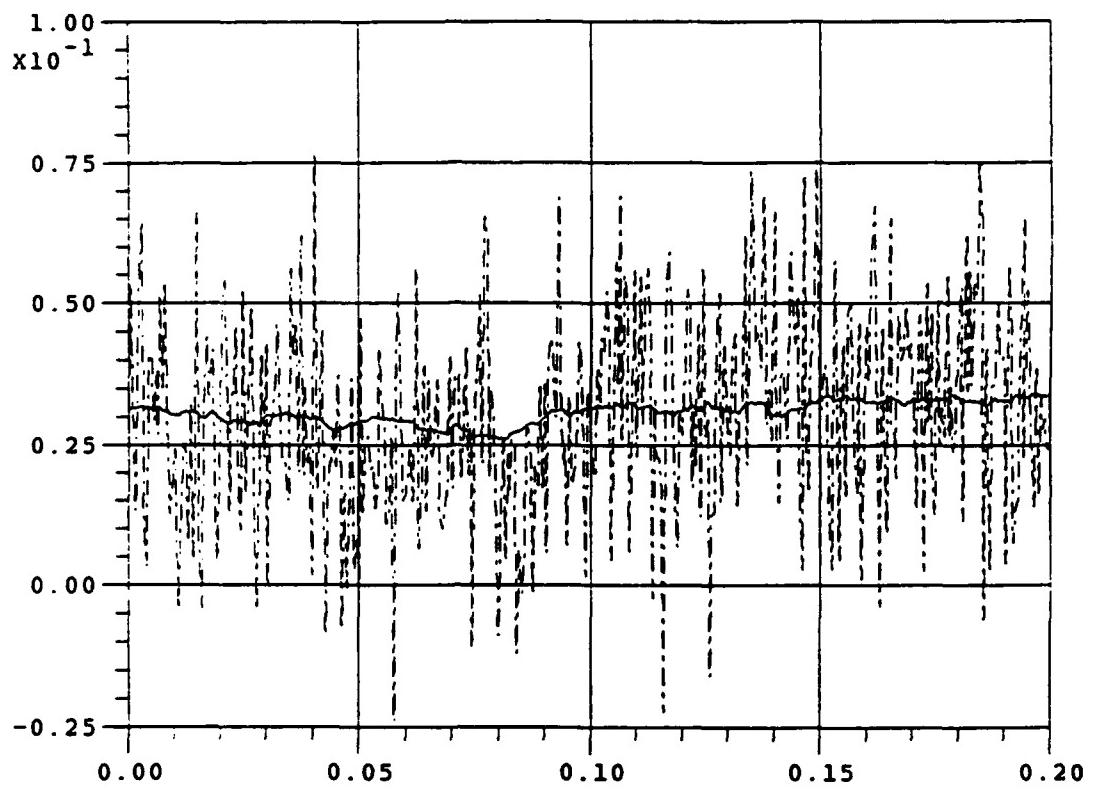
Figure 3.3: Lognormal sample path, SNR=5
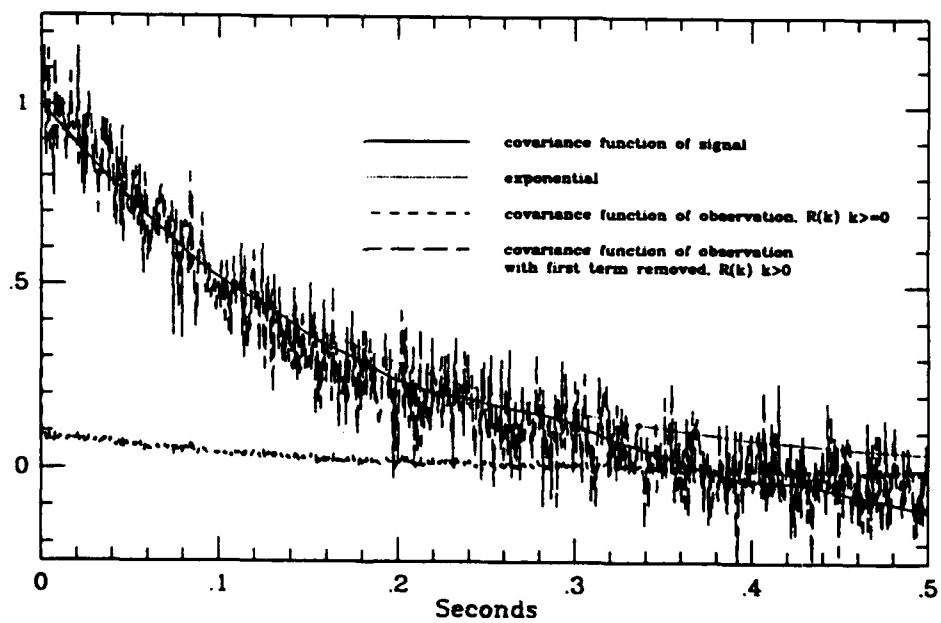
Figure 3.4: Covariance functions for Lognormal sample path, SNR=5
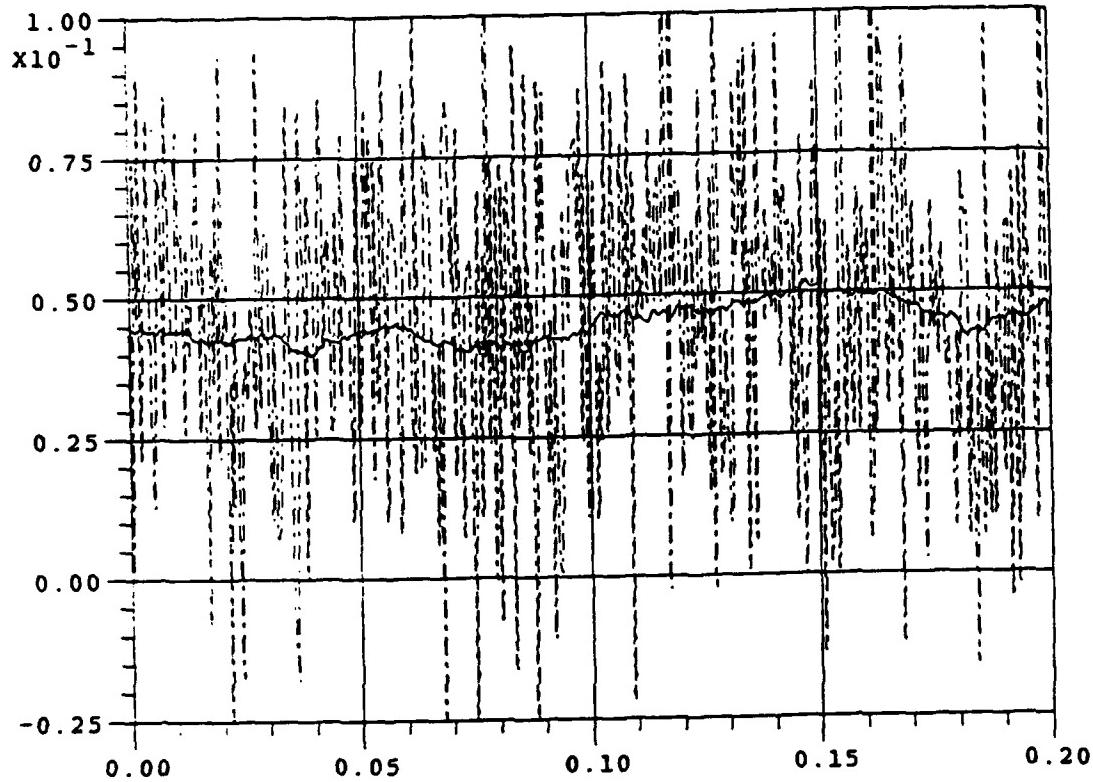
Figure 3.5: Lognormal sample path, SNR=2
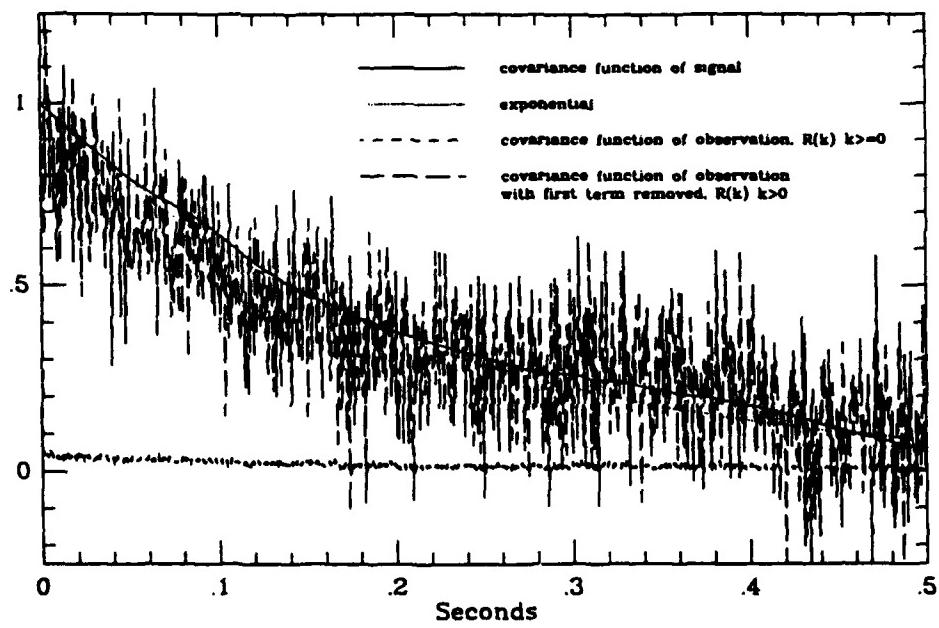


Figure 3.6: Covariance functions for Lognormal sample path, SNR=2

that the detector becomes numerically unstable for large SNR (ie. small additive noise power) this estimator was not implemented.

An alternate method of estimating the Rayleigh parameters is to estimate ρ_G from the covariance function of the output data similar to what we did in the lognormal case. If y_i is correlated Rayleigh (amplitude) data and we define $z_i = y_i^2$ then z_i is correlated Rayleigh power data. If the underlying state is Gaussian with exponential decorrelation and time constant t_G then z_i will have exponential decorrelation with time constant $t_z = \frac{t_G}{2}$ [Bar80]. Figures 3.7 to 3.9 show the normalized covariance functions of a Rayleigh amplitude signal, the square of the Rayleigh amplitude signal and a matched exponential function. We see that the decorrelation time $t_r \approx t_z$. We note that as the decorrelation time increases, $\rho_G \rightarrow 1$, and the covariance function does not go smoothly to zero. The oscillatory behavior is due to numerical instabilities in the discretization scheme as $\rho_G \rightarrow 1$ and is known as ringing. Figures 3.10, 3.12 and 3.14 are plots of Rayleigh (amplitude) model sample paths without noise and with additive noise for SNR 5. Figures 3.11, 3.13 and 3.15 are plots of the normalized covariance functions of the signal without noise, the signal plus noise $R_o(k), k \geq 0$, signal plus noise with the first term removed $R_o(k), k > 0$, and an exponential function. As noted earlier we have the desired behavior that the covariance function of the observation is approximately the sum of an exponential function, due to the signal, and a δ function due to the additive noise.

We can estimate $c\sigma_G$ from the ML estimator for $c^2\sigma_G^2$ for i.i.d. observations which is given by

$$c^2\sigma_G^2 = \frac{1}{2n}\sum_{i=1}^{n} y_i^2 \ .$$

Alternately, we could use the power to median or the power to mean ratios for i.i.d. observations given by (reference appendix C)

$$\frac{\text{power}}{\text{median}} = \frac{2}{\sqrt{\ln(4)}}c\sigma_G$$

$$\frac{\text{power}}{\text{mean}} = \sqrt{\frac{8}{\pi}} c \sigma_G \ .$$

3.4 Estimation of noise scaling γ

We will estimate γ from the covariance function of the observations. The noise in the output equation is white zero mean Gaussian so its covariance function is a δ function. Since the noise is independent of the signal $h(\mathbf{x})$ the covariance function of the observation data is the sum of the covariance functions of the noise and the signal. We use $R(k)$ to denote the covariance function, with subscripts s,n,o to denote signal, noise, and observation respectively. Under the model assumptions made the signal covariance function is approximately exponential so $R_s(0)$ is approximately $R_s(1)$ provided the decorrelation time constant of the signal is not on the order of the time discretization Δt. With the noise covariance function assumed to be a δ $R_s(1) \approx R_o(1)$. So we have

$$R_o(0) = R_s(0) + R_n(0)$$

or

$$R_n(0) = R_o(0) - R_s(0) \ .$$

So $R_s(0) \approx R_o(1)$ and substituting we get

$$R_n(0) \approx R_o(0) - R_o(1) \ .$$

In figures 3.11 to 3.15 as the decorrelation time approaches the time discretization step size Δt the δ function, due to the noise, becomes less distinct because the covariance function of the signal is approaching a δ function with respect to the time step size Δt. Additionally, if the signal to noise ratio is large, ie. the noise covariance function has a δ function with magnitude on the order of $R_s(0) - R_s(1)$, then $R_n(0) \approx R_o(0) - R_o(1)$ is clearly a poor approximation which could cause significant numerical problems and errors since the observations are
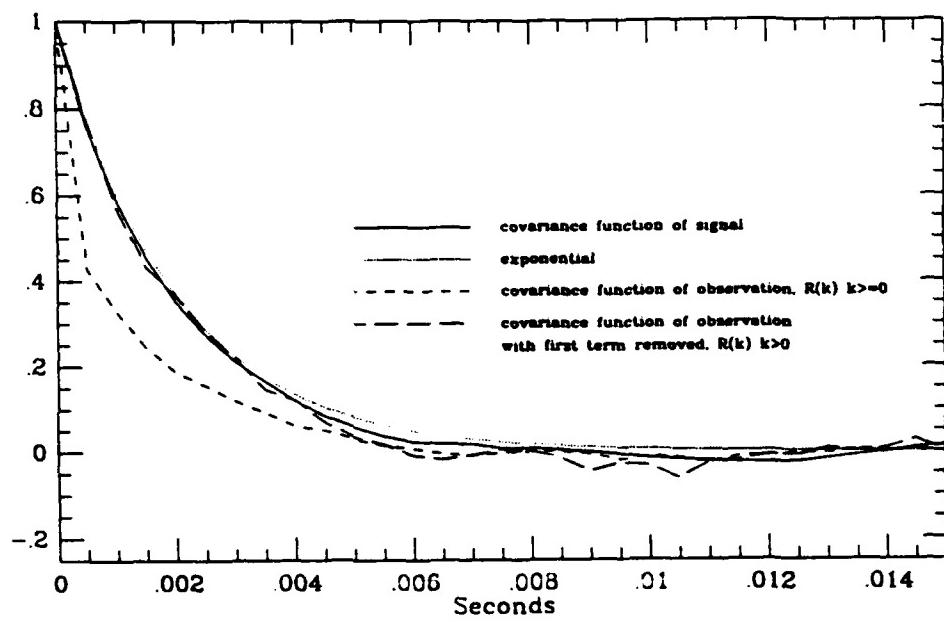
Figure 3.7: Covariance functions for Rayleigh amplitude and Rayleigh power sample path, $t_r \approx t_l/100$
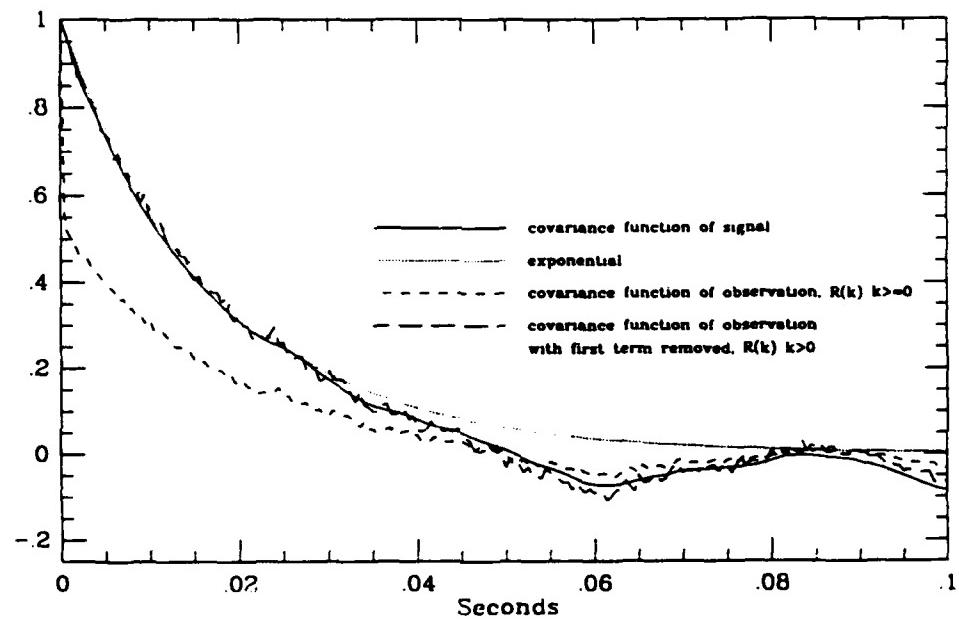
Figure 3.8: Covariance functions for Rayleigh amplitude and Rayleigh power sample path, $t_r \approx t_l/10$
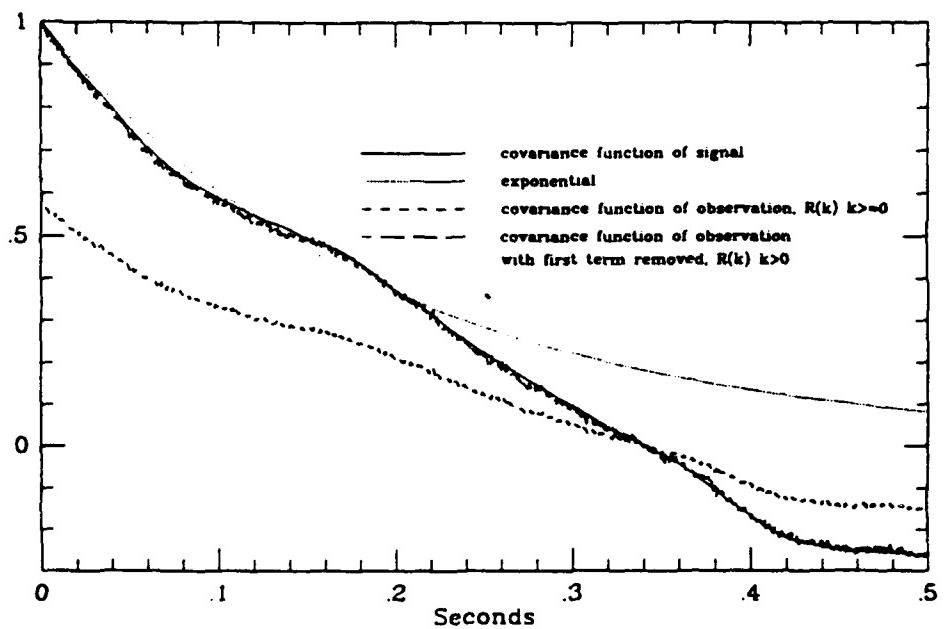
Figure 3.9: Covariance functions for Rayleigh amplitude and Rayleigh power sample path, $t_r \approx t_l$
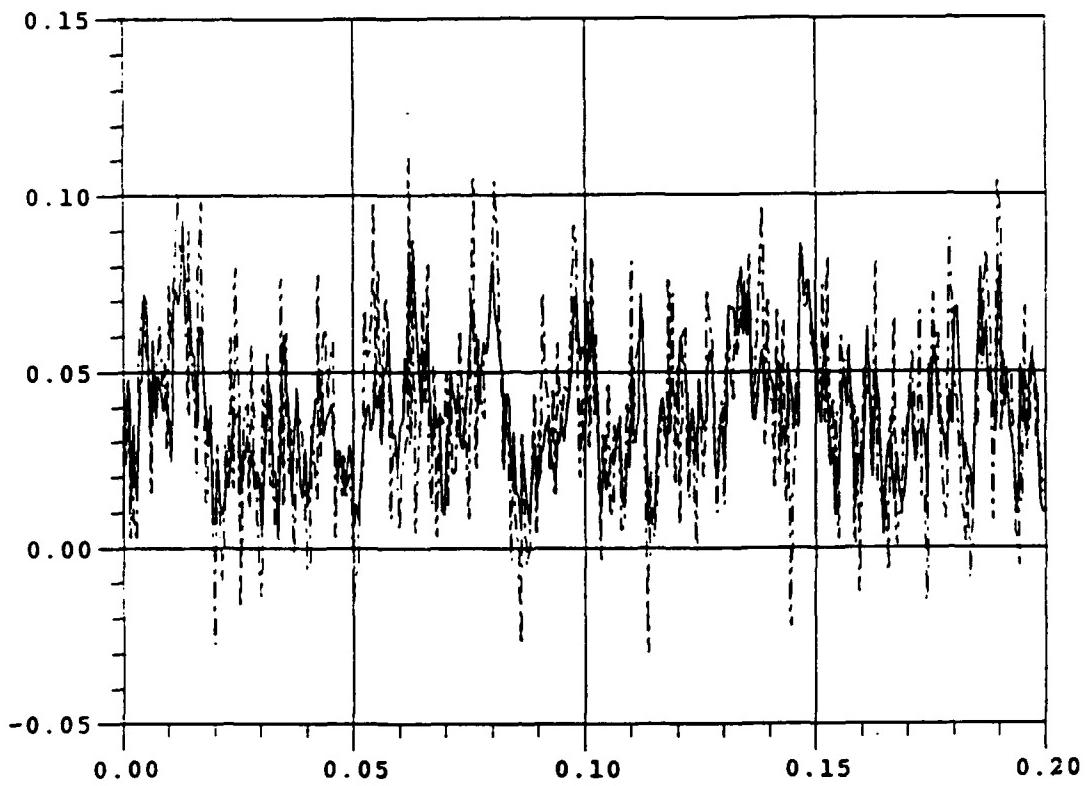
Figure 3.10: Rayleigh sample path,$t_r \approx t_l/100$ SNR=5
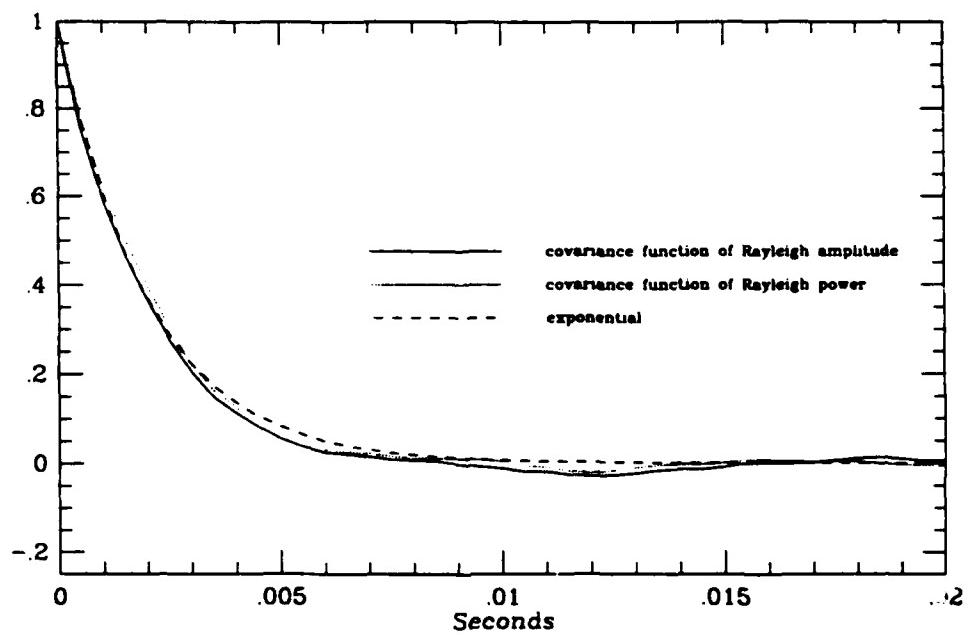
Figure 3.11: Covariance functions for Rayleigh sample path, SNR=5
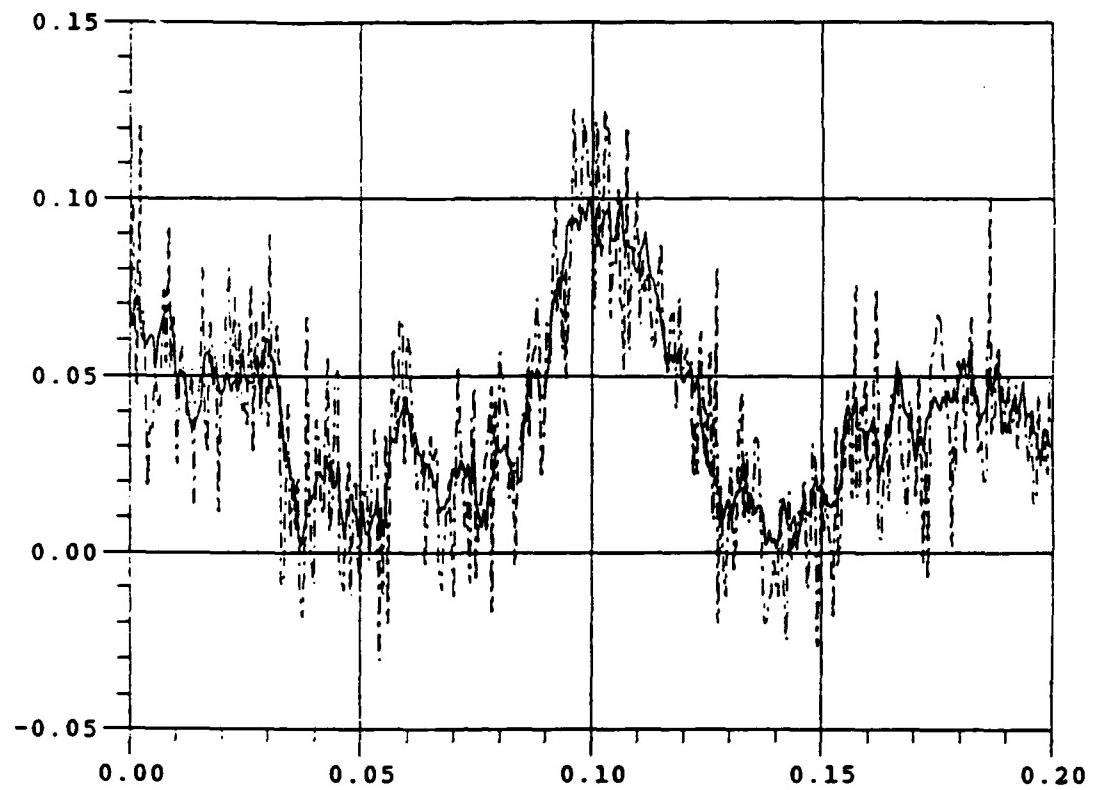
Figure 3.12: Rayleigh sample path,$t_r \approx t_l/10$ SNR=5

covariance function of Rayleigh amplitude
covariance function of Rayleigh power
exponential
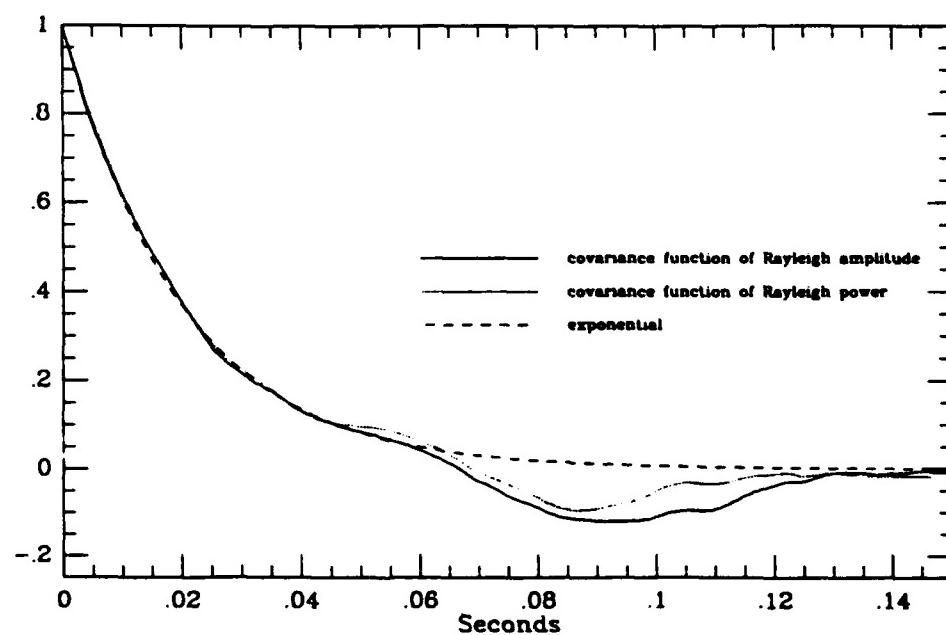
Figure 3.13: Covariance functions for Rayleigh sample path, SNR=5
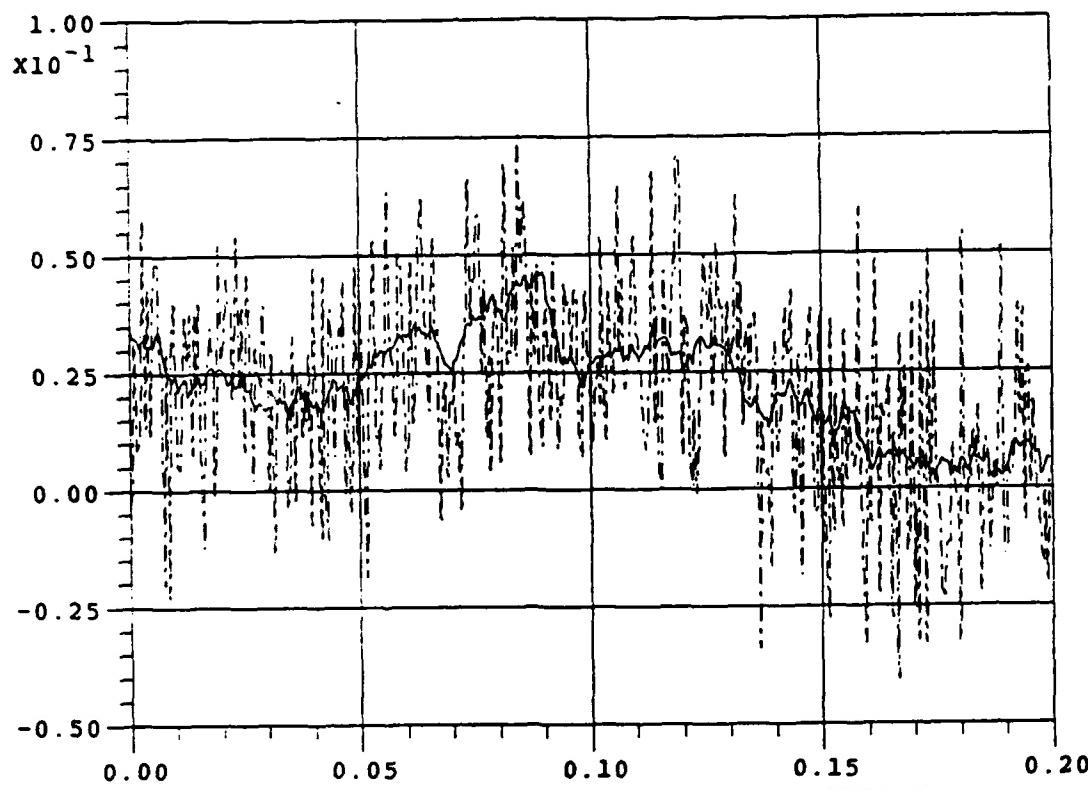
Figure 3.14: Rayleigh sample path,$t_r \approx t_l$ SNR=5
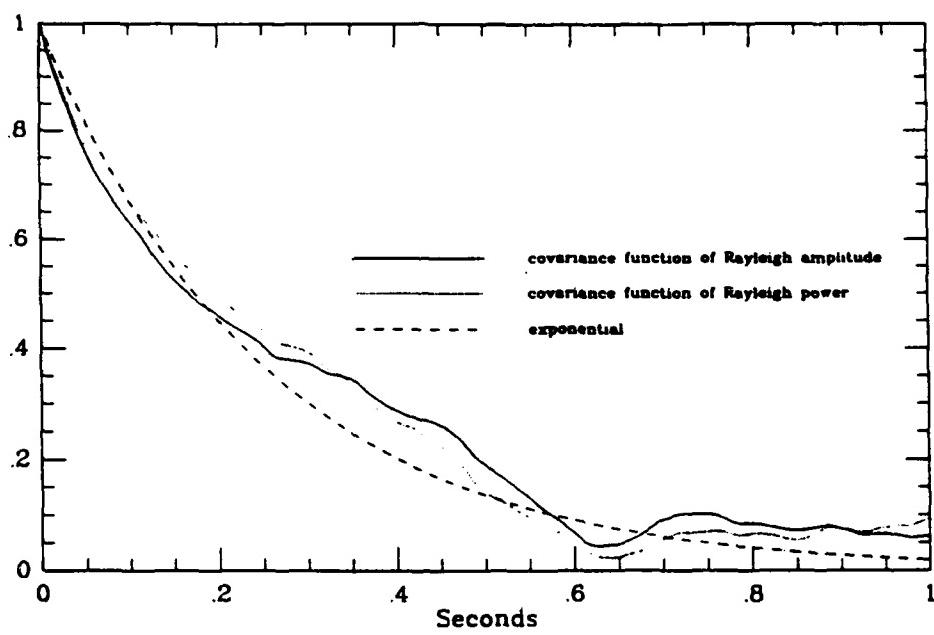
Figure 3.15: Covariance functions for Rayleigh sample path, SNR=5

renormalized by $\frac{1}{\gamma}$ where $R_n(0) = \frac{x^2}{\Delta t}$ after discretizing. Again, for the models we use, figures 3.11, 3.13 and 3.15 exhibit the desired behavior of a δ function plus an exponential function. This estimate could be improved by extrapolating back to $R_o(0)$ after matching the exponential to $R_o(k)$ $k > 0$. This would help remove variations of $R_o(k)$ from $R_s(k)$ for $k > 0$ due to the additive noise not being an exact δ function as seen most significantly in figures 3.2, 3.4, and 3.6.

Chapter 4

Results

4.1 Model specification

We choose the Rayleigh and lognormal models based on NRL data. Tables 4.1 and 4.2 summarize the parameter values chosen to test the detectors performance. Results from Chapter 3 and the appendices were used to determine parameter values and signal and noise statistics. The Rayleigh model with parameters $A = -24.06, B = 1.038$ and $C = .2117$ corresponds to the Chaff/decoy. We will denote the hypothesis test between this particular Rayleigh model and the lognormal model as the ship versus decoy case. The other sets of parameters are to test the effect of different characteristics of the signal and noise on the detector performance. The decorrelation time and the stationary mean are investigated. The second set of five parameters for the Rayleigh have the stationary mean shifted by 5 percent and the powers moved closer. In addition to

Lognormal parameters			stationary statistics				$\approx t_l$
Q	R	S	σ_G^2	μ_l	σ_l^2	power	
-4.800	.4646	.03935	.02251	.03980	.00003605	.001620	.2066

Table 4.1: Lognormal model parameters

Rayleigh parameters			stationary statistics				
A	B	C	σ_G^2	μ_r	σ_r^2	power	$\approx t_r$
-.04840	.04668	.2117	.02251	.03980	.0004330	.002018	$100tc_l$
-.2420	.1044	.2117	.02251	.03980	.0004330	.002018	$10tc_l$
-2.419	.3299	.2117	.02251	.03980	.0004330	.002018	tc_l
-24.06	1.038	.2117	.02251	.03980	.0004330	.002018	$tc_l/10$
-227.9	3.111	.2117	.02251	.03980	.0004330	.002018	$tc_l/100$
-.04840	.04668	.2011	.02251	.03781	.0004330	.001821	$100tc_l$
-.2420	.1044	.2011	.02251	.03781	.0004330	.001821	$10tc_l$
-2.419	.3299	.2011	.02251	.03781	.0004330	.001821	tc_l
-24.06	1.038	.2011	.02251	.03781	.0004330	.001821	$tc_l/10$
-227.9	3.111	.2011	.02251	.03781	.0004330	.001821	$tc_l/100$

Table 4.2: Rayleigh model parameters

looking at different values for a, b, c, q, r and s we also investigate the detectors behavior for different SNR. Table 4.3 summarizes the values of γ and the corresponding SNR. We also tested the detectors behavior for different values of P_F and P_M though for most simulations $P_F = P_M = .001$. We made a more indepth comparison of the implicit and the mixed schemes implementing (2.4) than it was done in Chapter 1. The probability of detection, the probability of

γ	power	SNR w.r.t stationary Lognormal power
.0006364	.000810	2
.0004025	.000324	5
.00020125	.000081	20

Table 4.3: Noise scale

miss and the average detection time are calculated from 1000 simulation runs.

4.2 Simulations

The following section summarizes results of simulations for the models specified in tables 4.1 and 4.2 for the SNR given in table 4.3. For these simulations $P_F = P_M = .001$. Effects of the decorrelation time, SNR, stationary mean and implicit versus mixed scheme on the percentage of correct decisions and average decision time are given. The results are based on 1000 runs. Figures 4.1 to 4.8 give results for the percentage of correct decisions versus the decorrelation time (of the Rayleigh relative to the lognormal) under hypothesis 0 and 1. The plots are for SNR of 20, 5 and 2, with the (stationary) means matched and shifted by 5 percent, and for the implicit and mixed scheme implementing (2.4). Figures 4.9 to 4.16 give the corresponding plots for the average detection time versus decorrelation time. These plots correspond to tables E.1 to E.6 in appendix E.

4.2.1 Effects of decorrelation time, SNR, mean, and scheme on percentage of correct decisions

Under hypothesis 1 (lognormal) the detector performs well with respect to the percentage of correct decisions except for the case $t_r \approx t_l/100$. In this case the detector performs best for SNR=20, and progressively worse for SNR=5 and 2. One possible reason for this behavior has already been mentioned in chapter 3; that is as the SNR decreases (more noise) the covariance function of the observation for the lognormal tends toward a δ function with the lognormal signal behaving like a mean for the noise and being removed. In the cases $t_r \approx t_l/100$ and $t_l/10$ the theoretical decorrelation time of the Rayleigh is approaching the time step size Δt and the covariance function of the signal is becoming a δ function. So we see that the covariance function of the lognormal plus noise is approaching the covariance function of the Rayleigh. This behavior is shown in

44
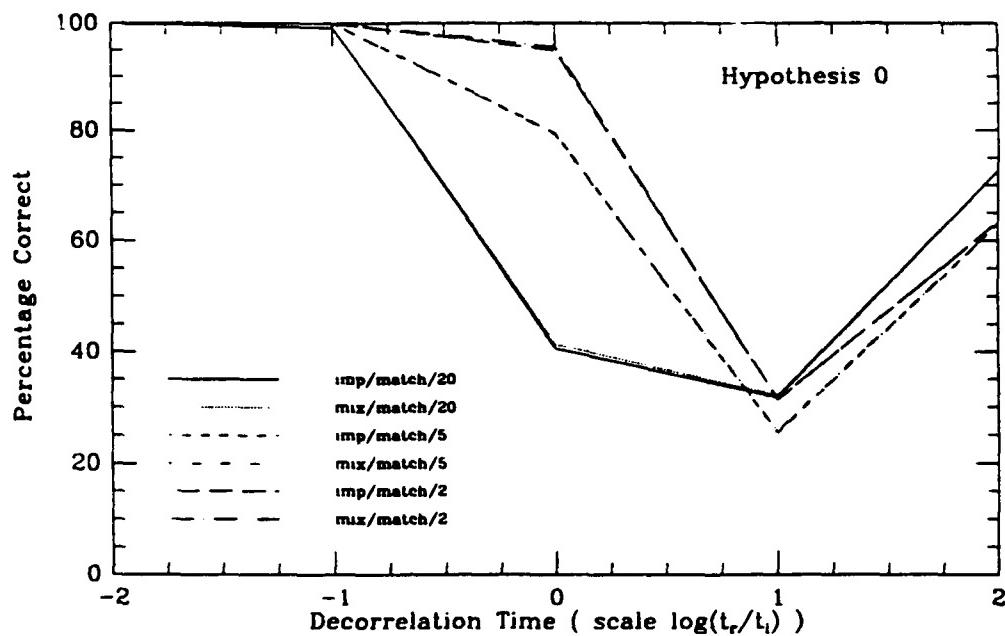
Figure 4.1: Percentage of correct decisions under Hypothesis 0
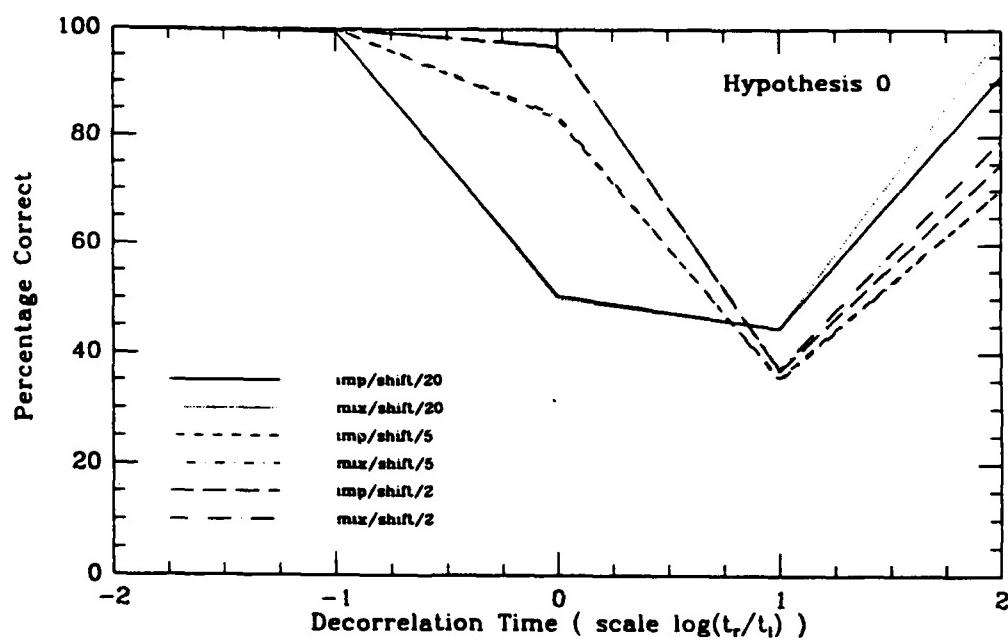
Figure 4.2: Percentage of correct decisions under Hypothesis 0

45
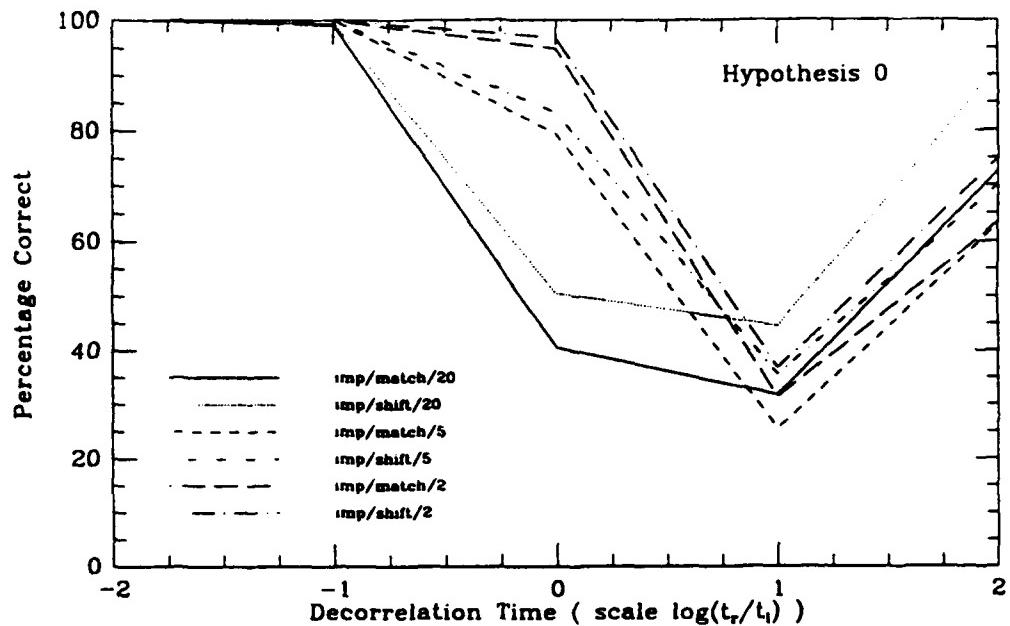
Figure 4.3: Percentage of correct decisions under Hypothesis 0
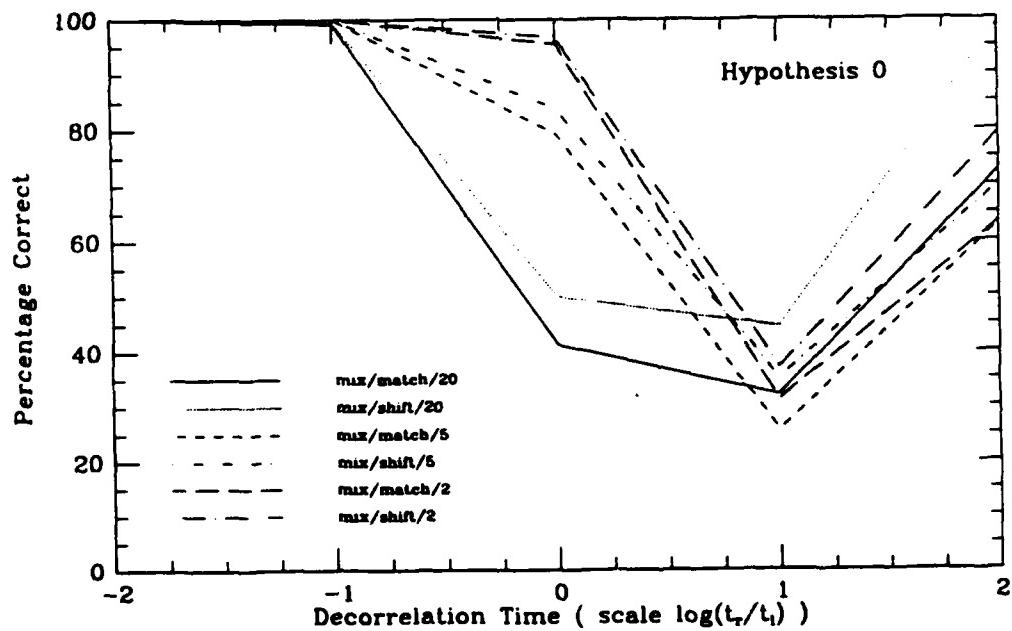
Figure 4.4: Percentage of correct decisions under Hypothesis 0
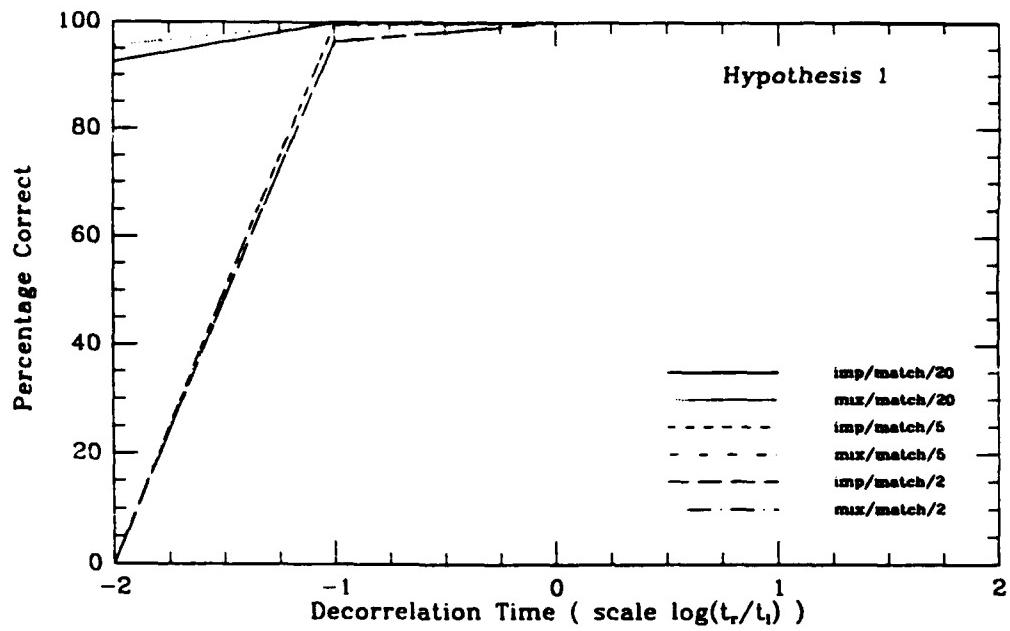
Figure 4.5: Percentage of correct decisions under Hypothesis 1
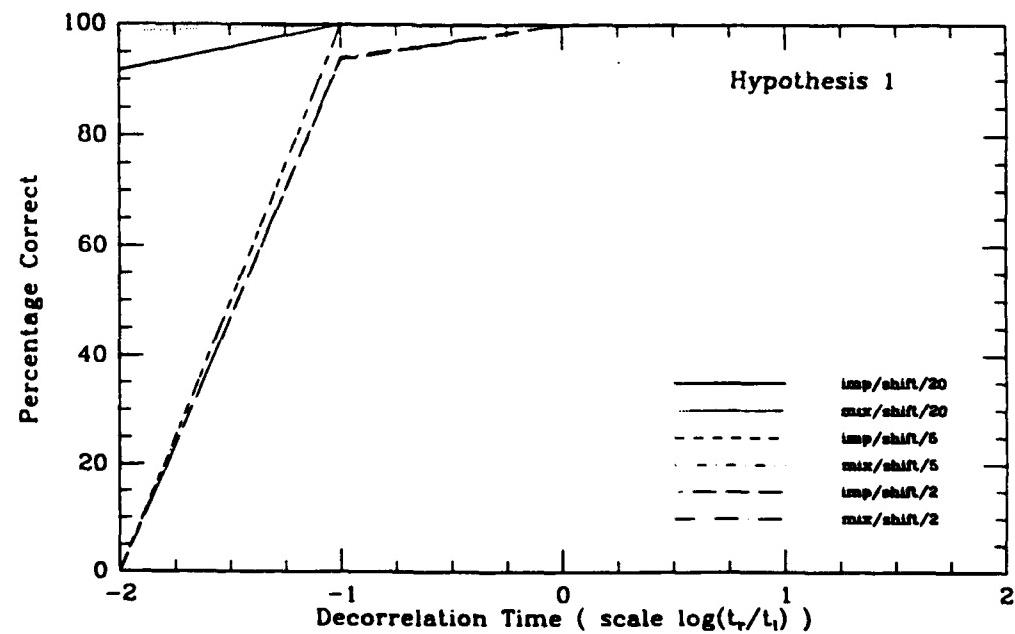

Figure 4.6: Percentage of correct decisions under Hypothesis 1
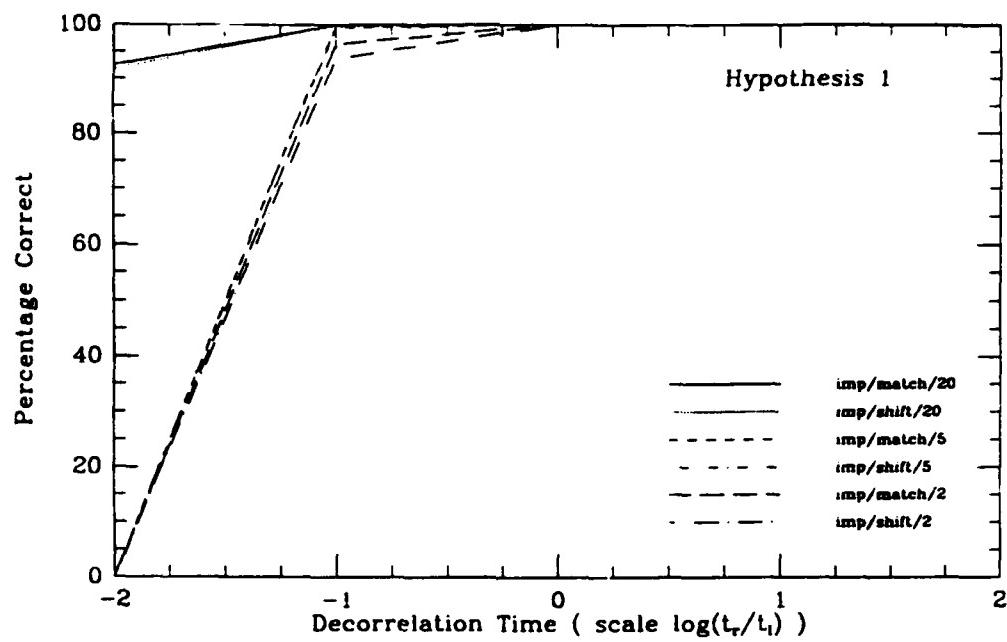
Figure 4.7: Percentage of correct decisions under Hypothesis 1
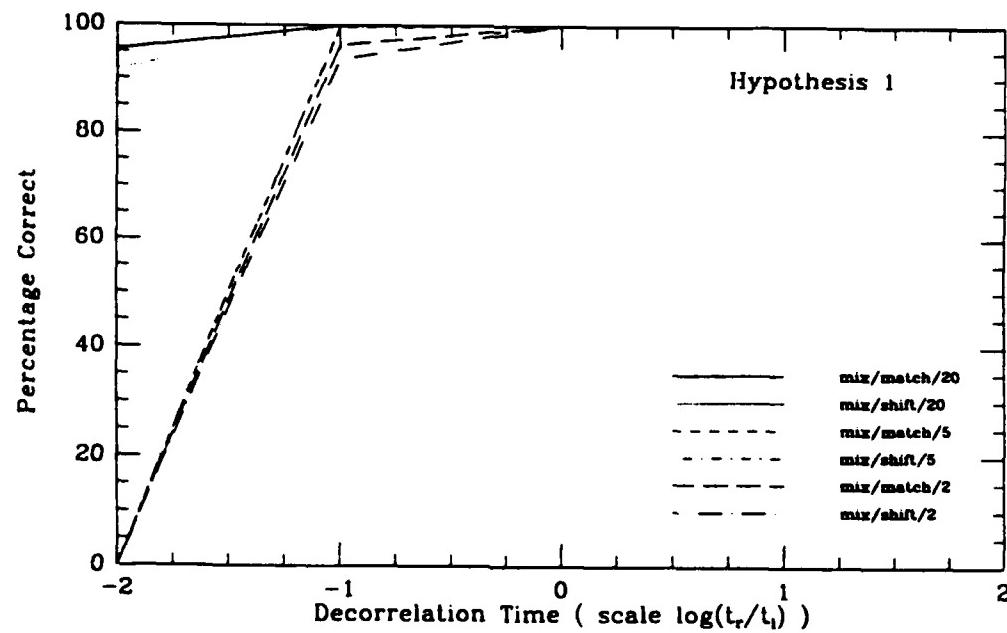
Figure 4.8: Percentage of correct decisions under Hypothesis 1

figures 3.1 to 3.6 and figures 3.10 to 3.15.

Under hypothesis 0, since this is a binary hypothesis test, as expected we see the detector works well for $t_r \approx t_l/100$ and $t_l/10$ and does poorly for $t_r \approx t_l.10t_l$ and $100t_l$. We also have the SNR effecting the percentage of correct decisions opposite its effect under hypothesis 1. This may be due to the fact that for a higher SNR (less noise, smaller γ) the observations, which are renormalized by $\frac{1}{\gamma}$, are weighted more in the solution to the Zakai equation and cause a quicker decision but for these longer decorrelation times fewer uncorrelated blocks of data are observed. The smaller SNR allows a longer observation.

Under both hypothesis we see that shifting the mean had only a small effect. Under hypothesis 1 there was a decrease in performance with respect to the percentage of correct decisions while under hypothesis 0 there was a slightly more significant increase in performance. Several runs were made when the means were significantly different (a factor of 10) and one could easily distinguish the two hypothesis. In this case the detector made correct decisions very quickly. However, when the means are so different these cases are not of interest since much simpler detectors could be implemented. With regards to the two scheme implemented, the implicit and mixed, for (2.4), we see that the percentage of correct decisions is nearly identical.

4.2.2 Effects of decorrelation time, SNR, mean, and scheme on average detection time

Under both hypothesis 1 and hypothesis 0 we see that for small SNR (more noise) the detector takes longer to make a decision which one would expect. This is due to the renormalization of the observations by $\frac{1}{\gamma}$. As γ increases (smaller SNR, more noise) the observations are weighted less in the solution to the Zakai equation.

We also see that the mean has a small effect on the average detection time. Under both hypothesis the shifted mean usually increase detection time slightly.

Figure 4.9: Average decision time under Hypothesis 0

Figure 4.10: Average decision time under Hypothesis 0
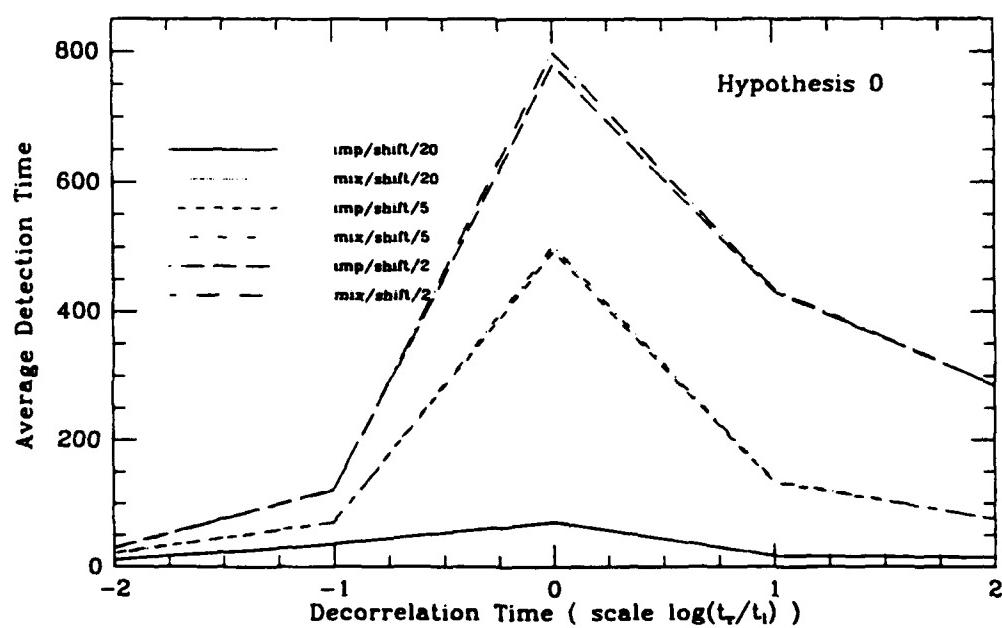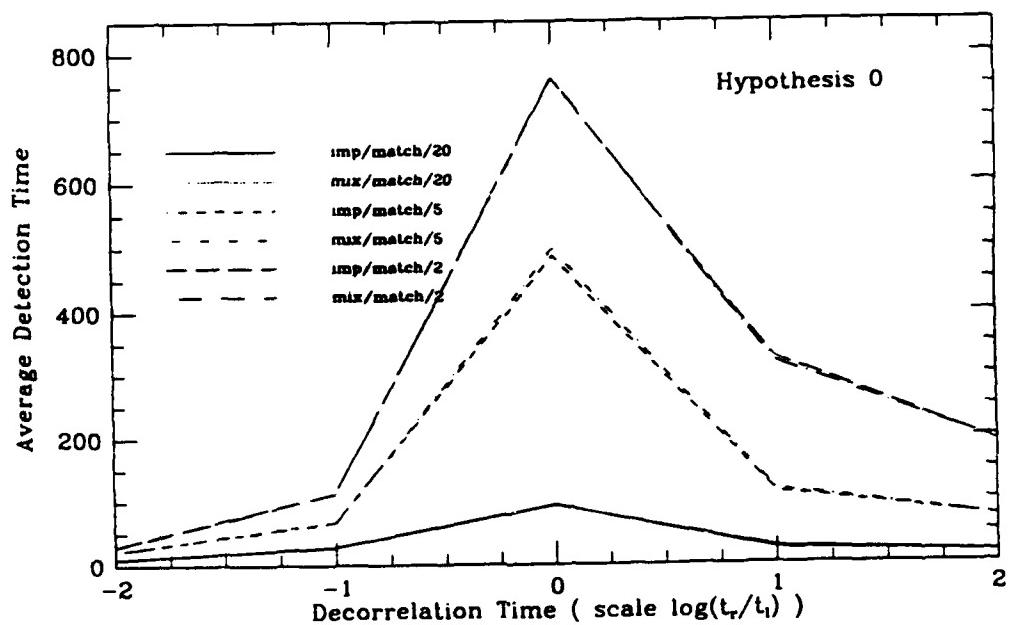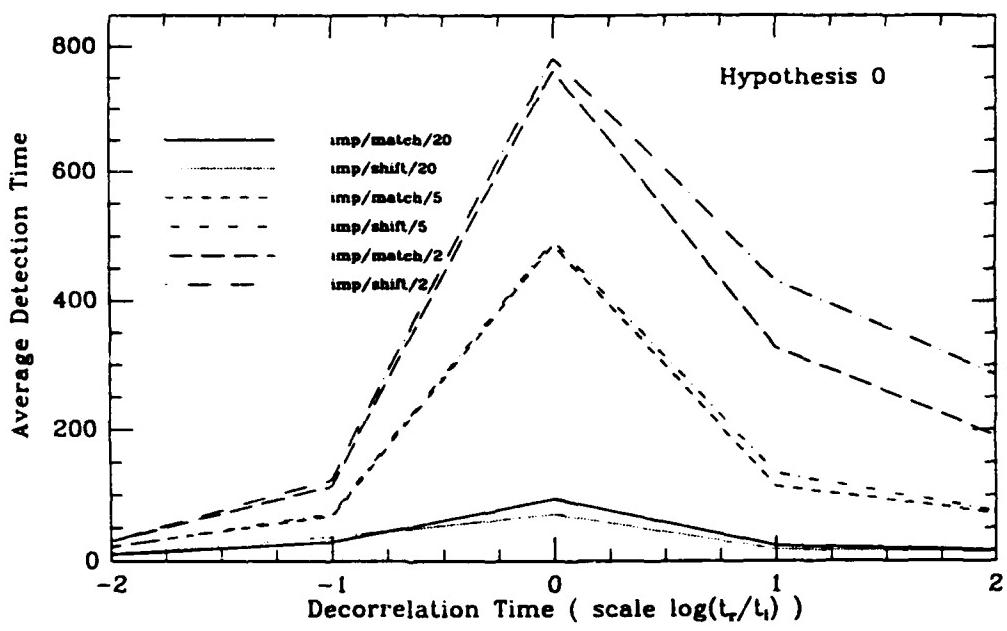
Figure 4.11: Average decision time under Hypothesis 0
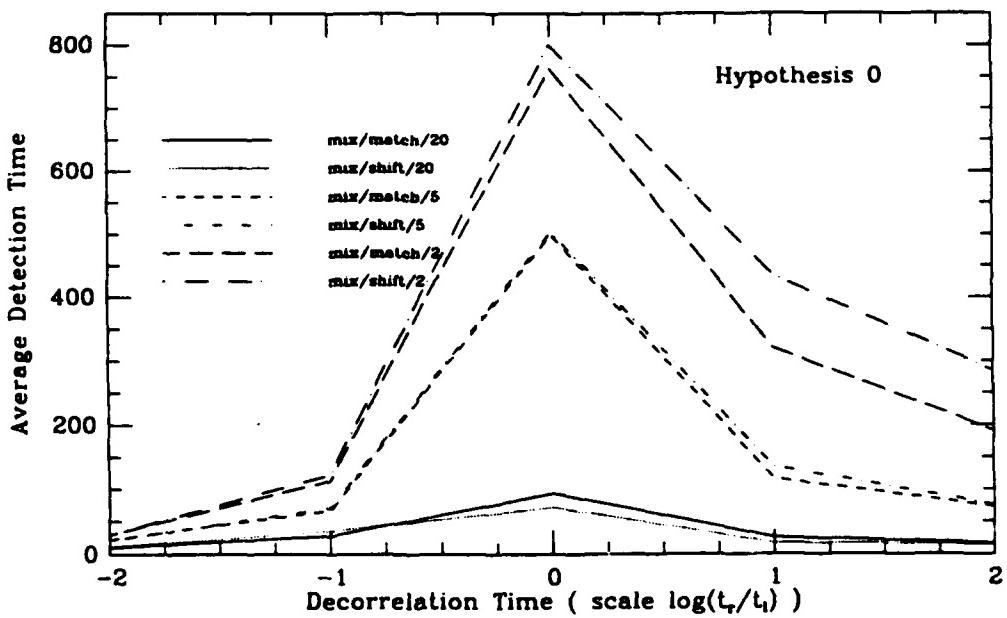
Figure 4.12: Average decision time under Hypothesis 0
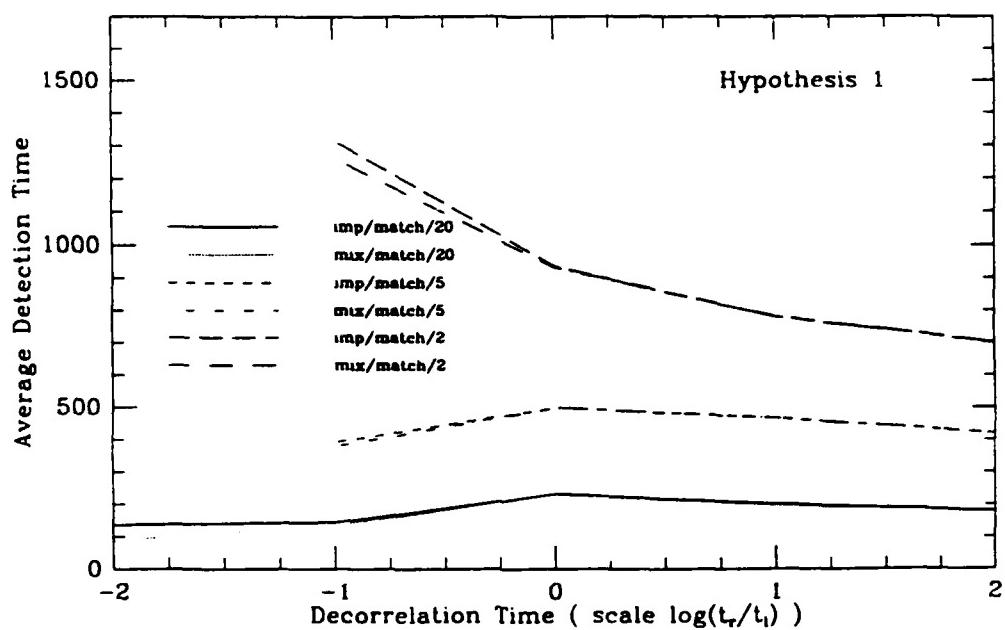
Figure 4.13: Average decision time under Hypothesis 1
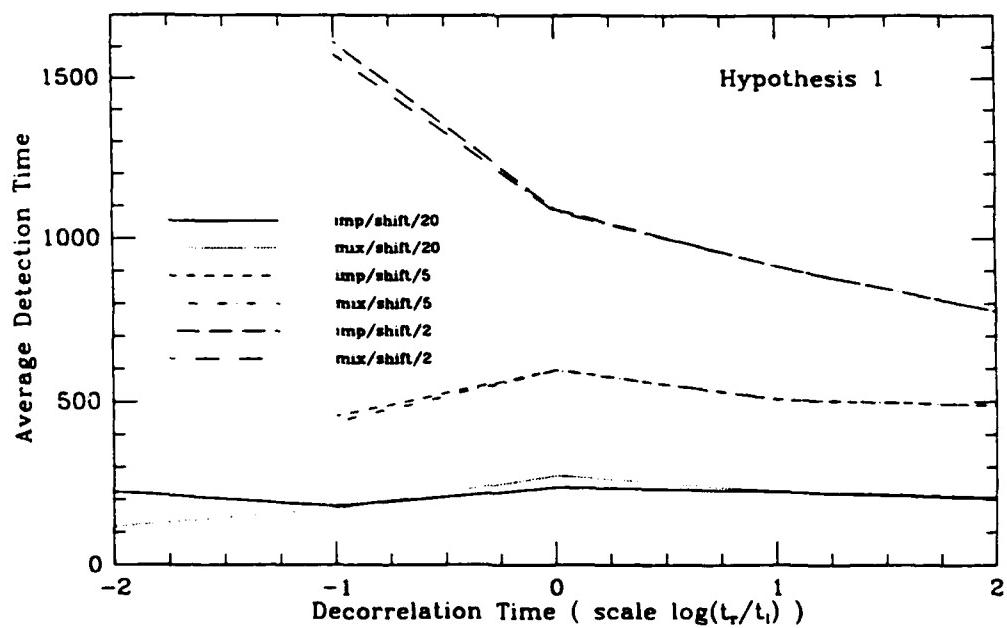
Figure 4.14: Average decision time under Hypothesis 1
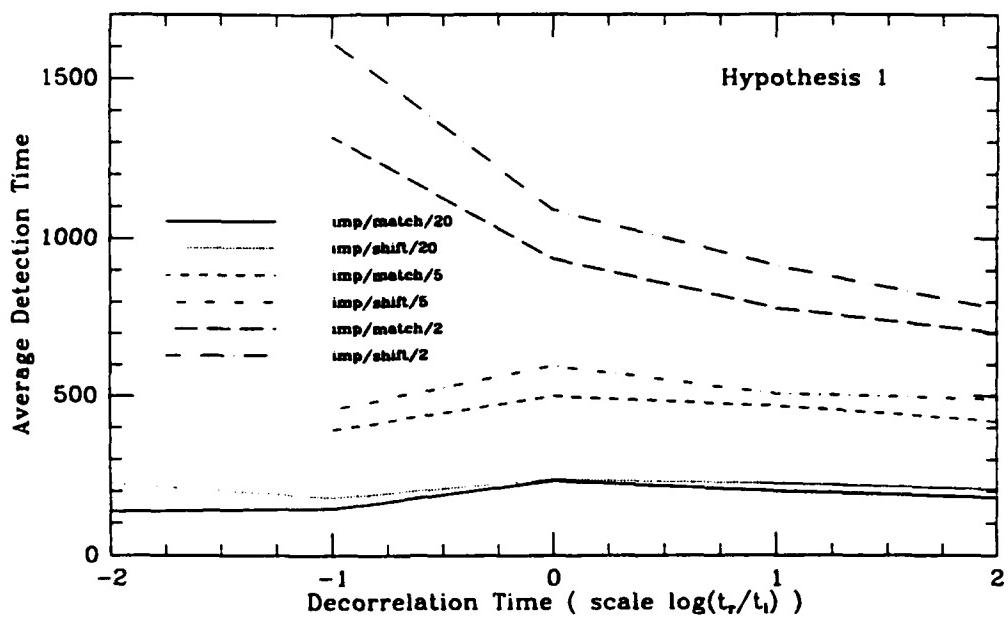
Figure 4.15: Average decision time under Hypothesis 1
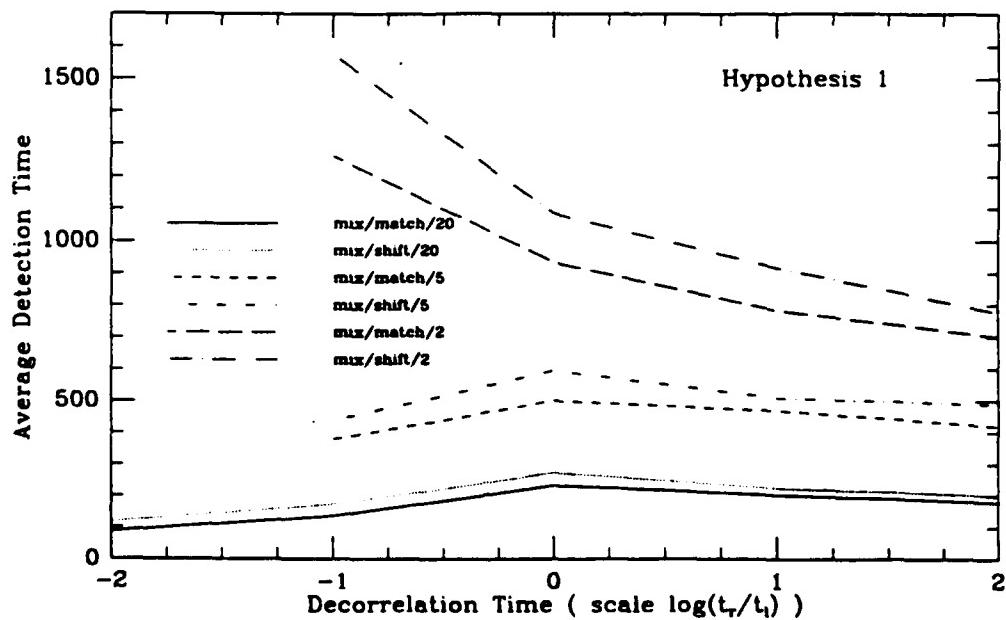
Figure 4.16: Average decision time under Hypothesis 1
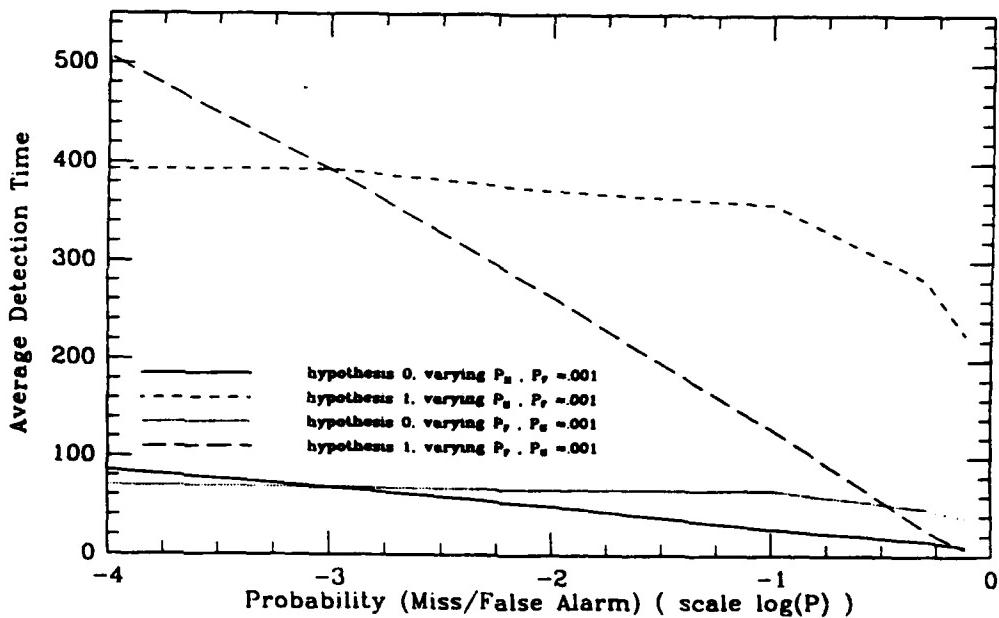
Figure 4.17: Average decision time for different values of P_F and P_M

Again, the implicit and mixed scheme give nearly identical results.

It is the decorrelation time that has the most significant effect on the average detection time. Under hypothesis 1 for $t_r \approx t_l/100$ we had no detections for SNR=5 and 2 so no average detection times are plotted. Except for hypothesis 1 with SNR=2 we see the detection time is peaked at $t_r \approx t_l$ and decreases as t_r increases or decreases. This implies the decorrelation time is a very significant feature of the signal for the detector. When the decorrelation times are matched the detector takes the longest to reach a decision. We also see that for the "symmetric" cases, ie. $t_r \approx t_l/100$ and $t_r \approx 100t_l$ or $t_r \approx t_l/10$ and $t_r \approx 10t_l$, that the detection time for the longer decorrelation times is longer.

4.3 Performance varying P_F and P_M

In figure 4.17 we have a plot of the average detection time for different values of P_F and P_M for the ship versus decoy case with the SNR=5. The figure corresponds to tables E.7 and E.8 in appendix E. We see that as the P_F or

| | Lognormal (Hypothesis 1) | | | |
	q	r	s	γ
actual	-4.80	.4646	.03935	.000403
estimated	-6.06	.4181	.03933	.000403

| | Rayleigh (Hypothesis 0) | | | |
	a	b	c	γ
actual	-24.06	1.038	.2117	.000403
estimated	-24.06	2.098	.1203	.000410

Table 4.4: Estimated parameters for the ship versus decoy case, SNR=5

the P_M increases, for fixed P_M and P_F respectivly, the detection time decreases. This is as one would expect since increasing P_F or P_M allows greater error. We see in tables E.7 and E.8 that the actual percentage of correct decisions is usually better than the theoretical values used to set the thresholds. For the cases where $P_F = .0001$ and $P_M = .0001$ the percentage of correct decisions is not significant since it is base on only 1000 runs, however, the average detection time still is.

4.4 Performance with estimated parameters

In chapter 3 we gave some justification for the models we chose to study the detector, as well as some techniques for estimating the parameters for these models. We applied some of these techniques to the ship versus decoy case and ran the detector with estimated parameters to determine what degradation would occur. For a better comparison of the results, the sample paths of the runs were the same as those used for the ship versus decoy case with SNR=5 and implicit scheme.

To estimate the parameters for the lognormal case we generated a correlated

	Hypothesis 0				
	% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
known	100.0	68.2	0	-	68.2
estimated	100.0	75.8	0	-	75.8

	Hypothesis 1				
	% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
known	99.5	393.0	0.5	164.4	391.9
estimated	99.6	294.1	0.4	159.5	293.5

Table 4.5: Comparison of ship versus decoy case with known and estimated parameters for the implicit scheme, SNR=5

sample path with SNR=5. We match an exponential function to the covariance function of the observation, $R_o(k)$ $k > 0$, and estimated the lognormal decorrelation time constant to be .16 seconds. We used the approximation

$$\frac{1}{t_G} = \left[\ln(\sigma_G^2) - \ln\left[\ln\left(1 + \frac{\exp(\sigma_G^2) - 1}{\exp(1)}\right)\right]\right]$$

to estimate the decorrelation time of the underlying Gaussian. We related the decorrelation time constant of the underlying gaussian to the diffusion parameter q using results from Appendix D to arrive at the estimate $q = -6.06$. From independent identically distributed data we used the median to estimate the scale parameter s to be .03933 . Using the mean to median ratio we estimated σ_G^2 to be .01444 whcih we related to the parameter r, using results in appendix D, to get the estimate .4181 for r.

To estimate the Rayleigh parameters we again generate a correlated sample path with SNR=5. We matched an exponential function to the covariance function of the observation, $R_o(k)$ $k > 0$, and estimated the Rayleigh decorrelation time constant to be .019 seconds. So the underlying Gaussian has an approx-

imate decorrelation time constant of .038 seconds. Using appendix D we got an estimate of -24.056 for a. From independent identically distributed data we used the average of the power to mean ratio which yielded an estimated $c\sigma_G$ of .03773 and the power to median ratio which resulted in an estimate of .03645 to arrive at the estimate $c\sigma_G = .03709$. As noted in chapter 3 since c and σ_G always appear together in the density we can only estimate there product so we arbitrarily set σ_G^2 for the Rayleigh model equal to that of the lognormal model. Thus using appendix D we estimated b to be 2.098 and the scale parameter c to be .1202. To get a better idea of the performance of the estimates for the Rayleigh model if we set the scale parameter c to its actual value of .2117 then the estimated value of b becomes 1.2122 which appears to be a much better estimate. Clearly, since $c\sigma_G$ appears together in the correlated Rayleigh density the signal statistics are unchanged regardless of how one assigns the values of c and σ_G provided their product remains the same. Clearly the conditional density $u(\mathbf{x}, t)$ which is the solution to the Zakai equation would be different though the statistics of the signal are theoretically the same. However, due to the number of different approximations used to implement a numerical solution it may be there is an advantageous way of assigning c and σ_G so as to minimize numerical errors. The effects of different values of c and σ were not studied in this thesis.

To estimate the noise scaling parameter γ we used the approximation

$$R_o(0) - R_o(1) \approx \frac{\gamma^2}{\Delta t} .$$

From the lognormal sample path γ was estimated to be .000403 and from the Rayleigh sample path .000410 which we averaged (.000407) to estimate γ.

We see from table 4.5 that the percentage of correct decisions are nearly identical for the detector with correct parameter values and for the detector with estimated parameter values. Under hypothesis 1 there is a slight increase for the average detection time of 11 percent while under hypothesis 0 there is a decrease in detection time of 25 percent for the estimated parameter values compared with the solution with correct parameter values.

4.5 Additional observations of the Zakai detector

In addition to the results already presented, we give some observations made in implementing this detector which one should be aware of. We have already commented on the fact that one must be careful in the choice of the time discretization step size Δt but one must also take care in the choice of the space (state) discretization. Numerical instabilities can occur for explicit schemes if the state discretization step size is to small relative to the time discetization size. Additionally, the computational complexity increases as the space discretization becomes more fine. However, if the discretization is too coarse then the discretized density cannot adequately represent the actual conditional density. As well as the necessity for appropriate choices of the space discretization and time discretization, the range over which the density is discretized must be chosen. If the range is too small the density will be truncated. If the range is too large then for a reasonable number of points in the discretization the density will again be too coarse. In figure 4.18 we have the plot of a signal and the conditional expectation of the signal generated from a discretization that was to course. One can see that the conditional density could not adequately represent the conditional density and the conditional expectation of the signal tended to have distinct values around which it fluctuated. In figure 4.19 we see the actual signal and the conditional expectation of the signal with a density discretized over too small a range. Note how the conditional expectation of the signal cannot track the signal beyond a certain point. This is because the discretization range truncated the density, not allowing nonzero values for the conditional density outside of the discretized range. The initialization of the underlying Gaussian state densities is another factor one should be aware of. In our simulations the densities are set to Gaussian distributions using the stationary variance and mean (which is zero). However, the actual sample path
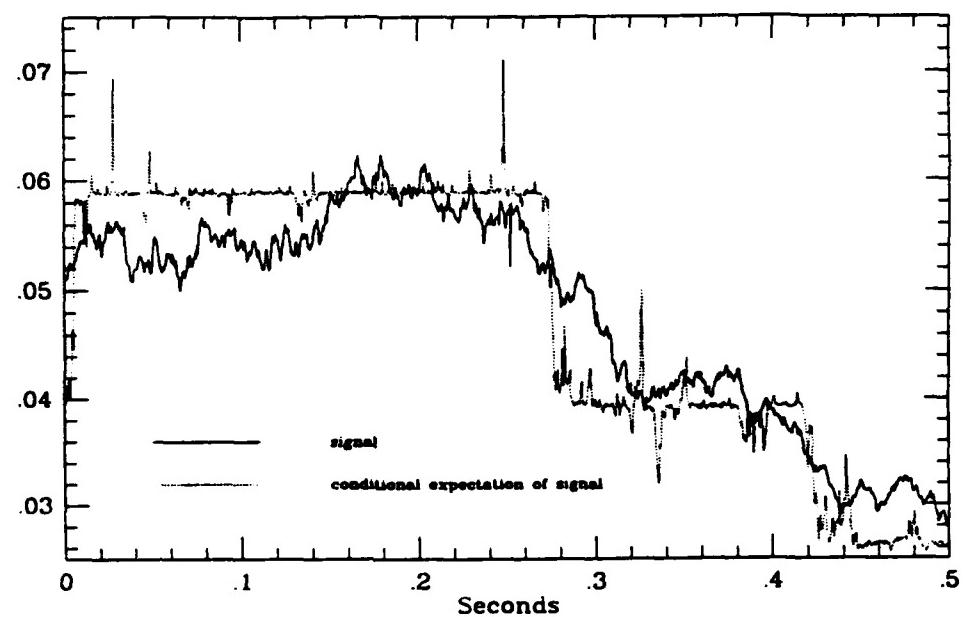
Figure 4.18: Signal and conditional expectation of the signal for to course a discretization
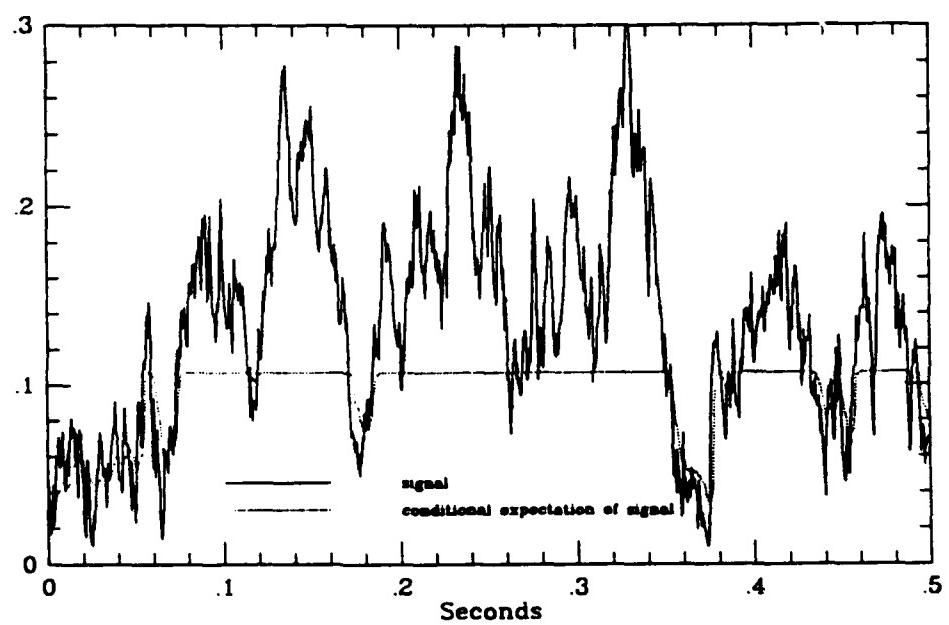
Figure 4.19: Signal and conditional expectation of the signal for to small a discretization range
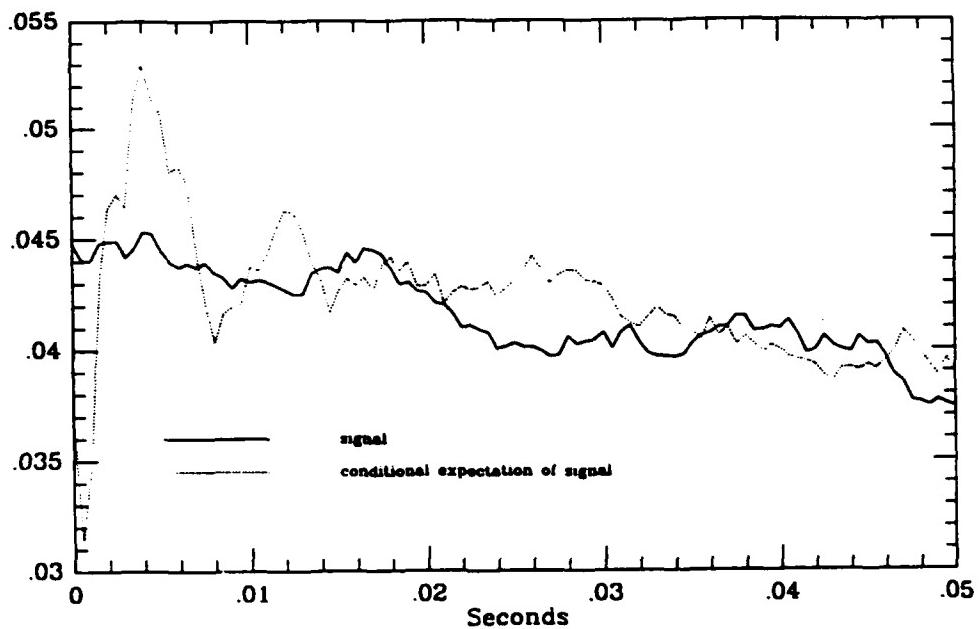
Figure 4.20: Signal and conditional expectation of the signal
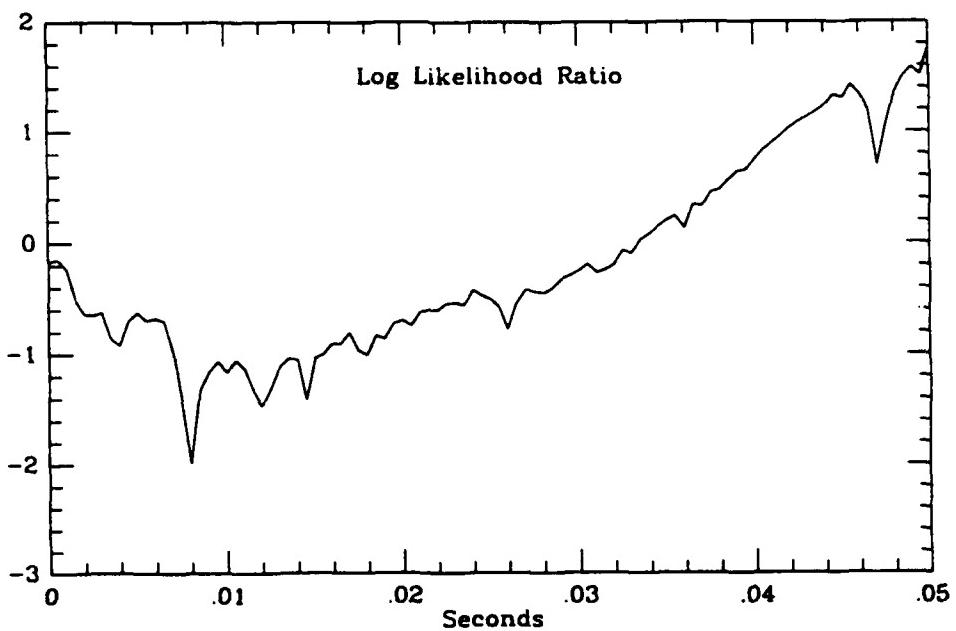
Figure 4.21: Log likelihood ratio for incorrect initialization of density

need not begin at the expected value and it may take several iterations before the conditional expectation of the signal yields a conditional expectation of the signal which approximates the actual signal as seen in figure 4.20. This can also cause initial errors in the log likelihood ratio as seen in figure 4.21 , where hypothesis 1 was initialized incorrectly. We see a general downward trend for .01 seconds (20 sample points) which corresponds to the approximate time it takes for the conditional density to yield a conditional expectation of the signal which tracks the signal.

4.6   The signal estimation problem

The detection and estimation problem are often related. In this case we are solving the Zakai equation for the unnormalized conditional density under two hypothesis conditioned on observation data. From the densities we can generate the log likelihood ratio and implement the sequential Wald formulation which is optimal in the previously defined sense. However, given the conditional densities an obvious estimator for the signal is the conditional expectation of the signal conditioned on the observations. In figure 4.22 to 4.25 we give the conditional expectation of the Rayleigh and lognormal signals under both hypothesis for the implicit scheme and SNR=5. We see that the Rayleigh model, which has a larger stationary variance, is better able to follow the more rapid variations of the signal while the lognormal model is slower and has more of a smoothing effect. We see that the solution of the Zakai equation not only yields a sequential detection scheme but a signal estimation scheme as well. We do not, however, make any evaluation of the perform   e of this estimator here.
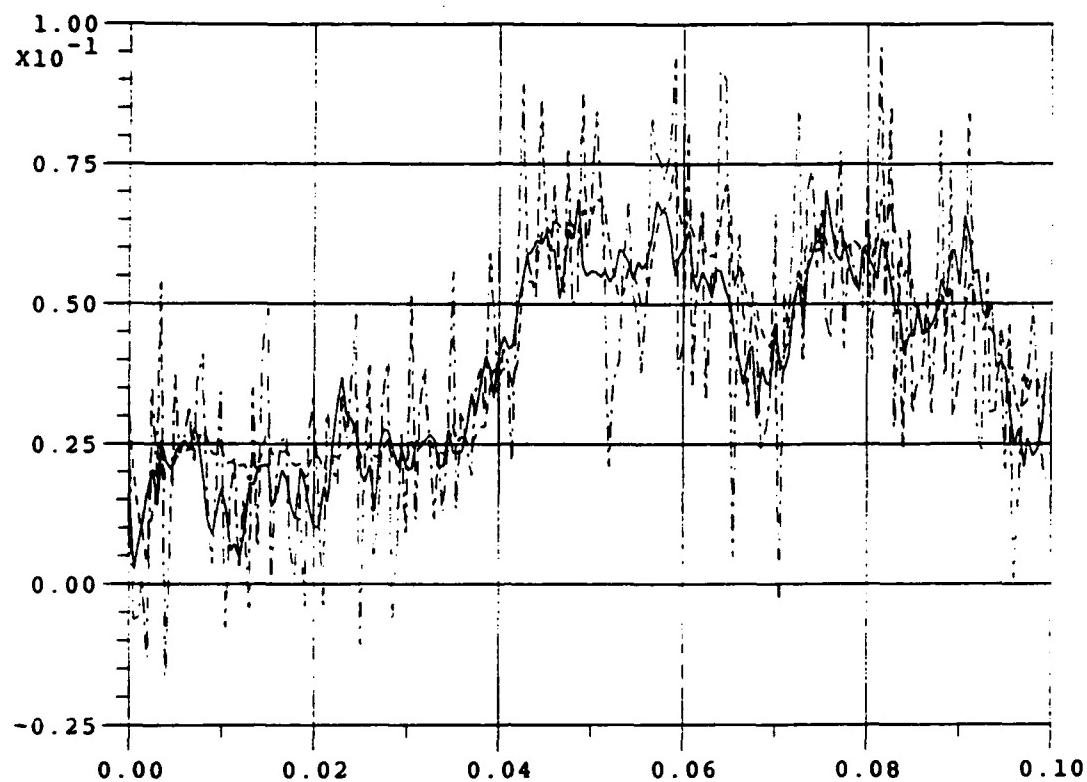
Figure 4.22: Expectation of Rayleigh signal given the Rayleigh, SNR=5
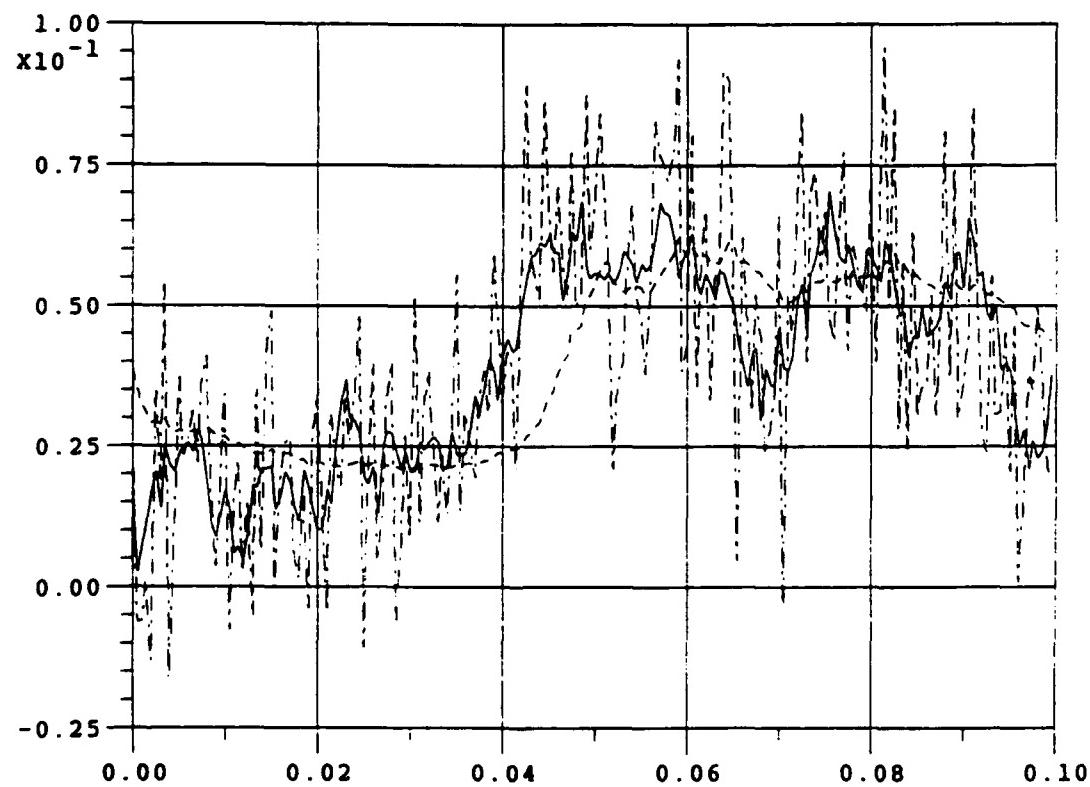
Figure 4.23: Expectation of lognormal signal given the Rayleigh, SNR=5
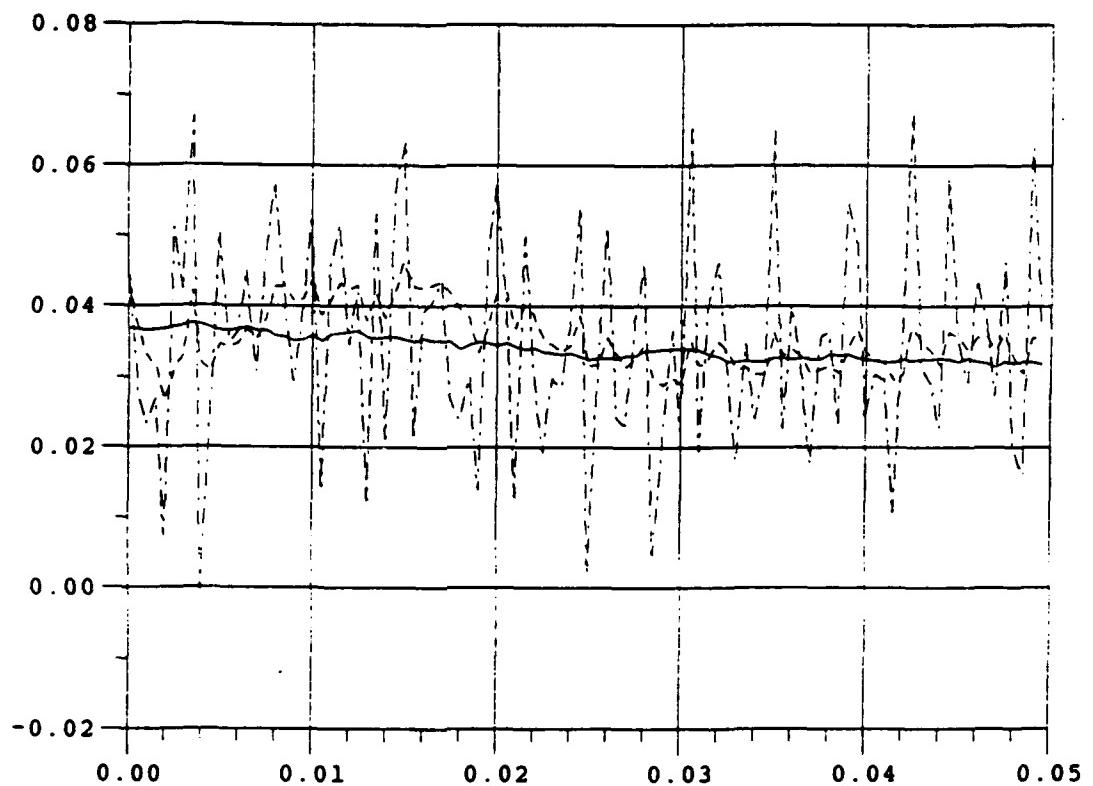
Figure 4.24: Expectation of Rayleigh signal given the lognormal, SNR=5
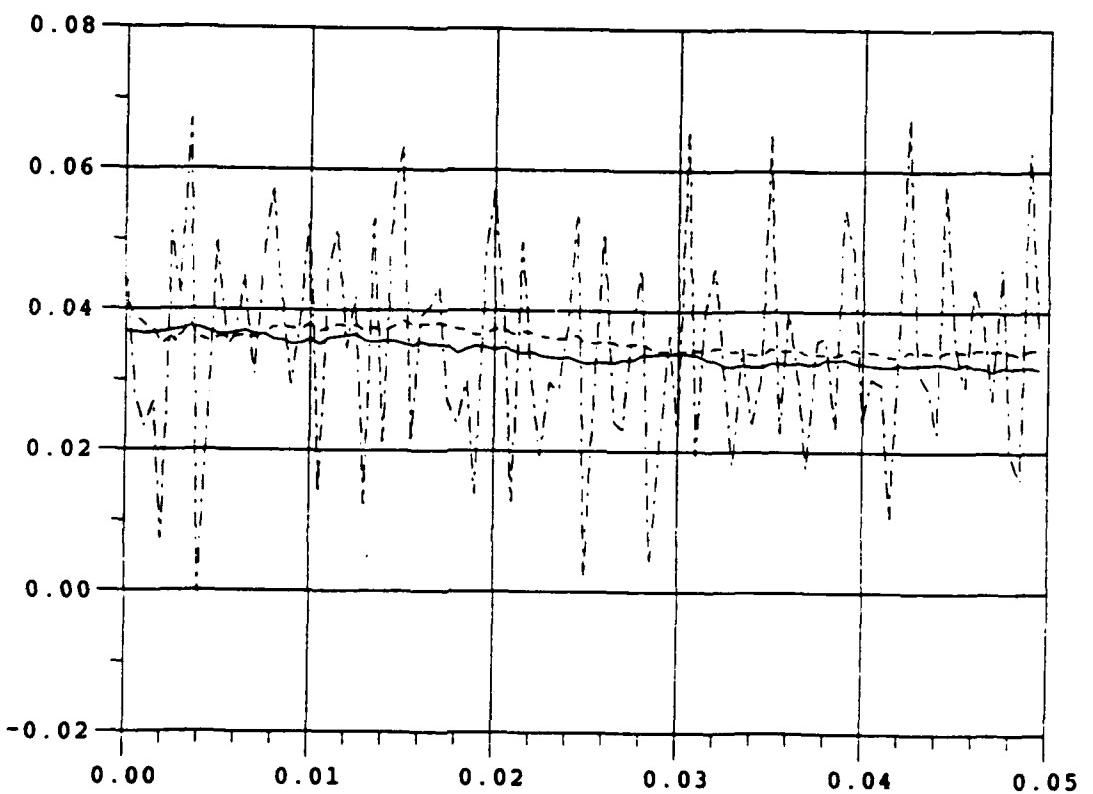
Figure 4.25: Expectation of lognormal signal given the lognormal, SNR=5

63

Chapter 5

Conclusions

We have presented an "optimal" sequential detector for the Wald formulation for diffusion processes for binary hypothesis testing. We gave several numerical schemes for implementing the detector for underlying states of dimension one and two which were studied to varying degrees. Different characteristics of the signal and discretization scheme were studied to determine their effects on the detector performance. It is clear that the performance varies significantly, depending on many factors including the relative decorrelation time of the signals, the SNR, the time discretization, and space discretization though not exclusively these factors. Because of the significant range of performance and the numerous parameters that gave rise to this range of performance, it is obvious that one must be an "expert" in order to make the detector useful. Also, the diffusion model represents a very large class of signals beyond the lognormal and Rayleigh models studied here for which the detector has not been tested. Additionally. the effect of many features of the signal where only partially studied. Clearly, much more work could be done in evaluating the detector's performance. In order to get a scheme that is computationally reasonable and performs well , which may not be possible, there will be a need for a number of simulations with different parameters. These facts point to a need for automating the generation of the detector with "expert" supervision for those not familiar in the

theory behind the Zakai detector. The MACSYMA code is an initial step in automating the process. The software system DEsign Laboratory for Processing Hidden Information (DELPHI) is under development which combines an AI engine, symbolic algebra, and numerical schemes as an "expert" system that not only generates and implements the Zakai detector but other detection schemes and model verification capabilities as well.

We also noted that the Zakai detector for hypothesis testing suggests a signal estimation scheme, namely the conditional expectation of the signal conditioned on the observations. We presented some results pertaining to signal estimation, however, no performance measure or comparison to other estimation schemes was performed.

ACKNOWLEDGEMENT

Appendix A

This appendix contails derivation of correlated lognormal and correlated Rayleigh densities.

Consider the vector

$$\left[\begin{array}{cccc} [x_{11} & x_{12} & \cdots & x_{1m}] \end{array} \begin{array}{cccc} [x_{21} & x_{22} & \cdots & x_{2m}] \end{array} \cdots \begin{array}{cccc} [x_{n1} & x_{n2} & \cdots & x_{nm}] \end{array} \right]$$

where x_{ij} are independent with respect to j and have exponential decorrelation with respect to i ie. $\mathbf{E}\,[x_{ij}x_{i+k,j}] = \mathrm{R}\,(k) = \sigma^2\rho^k$. Then the covariance matrix is symmetric with constant diagonals and $m - 1$ zero diagonals between each nonzero one. The covariance matrix is given by

$$Q = \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 & \rho & 0 & \cdots & 0 & \cdots\cdots & \rho^{n-1} \\ & & & & & & & & & \vdots \\ & & & & & & & & & \vdots \\ & & & & & & & & & 0 \\ & & & & & & & & & \vdots \\ & & & & \ddots & & & & & 0 \\ & & & & & & & & & \rho \\ & & & & & & & & & 0 \\ & & & & & & & & & \vdots \\ & & & & & & & & & 0 \\ & & & & & & & & & 1 \end{bmatrix}$$

Then Q^{-1} is symmetric and tridiagonal with $m - 1$ zero diagonals between the

main diagonal and the off diagonals and $m - 1$ ones on each end of the main diagonal. and is given by

$$Q^{-1} = \frac{1}{\sigma^2(1 - \rho^2)} \begin{bmatrix} 1 & & & -\rho & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ -\rho & & & 1+\rho^2 & & & & \\ & & & & \ddots & & & \\ & & & & & 1+\rho^2 & & -\rho \\ & & & & & & 1 & \\ & & & & & & & \ddots \\ & & & & & -\rho & & 1 \end{bmatrix}$$

This result was obtained using MACSYMA, More general results are given in [GL89] Finally, the determinant of Q is

$$|Q| = \sigma^{2nm}(1 - \rho^2)^{m(n-1)} \quad .$$

So for the lognormal case we have

$$Q_l = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ & \ddots & & \vdots \\ & & \ddots & \rho \\ & & & 1 \end{bmatrix}$$

$$Q_l^{-1} = \frac{1}{\sigma^2(1 - \rho)} \begin{bmatrix} 1 & -\rho & & \\ -\rho & 1+\rho^2 & \ddots & \\ & \ddots & \ddots & \\ & & 1+\rho^2 & -\rho \\ & & -\rho & 1 \end{bmatrix}$$

$$|Q_l| = \sigma^{2n}(1 - \rho^2)^{(n-1)}$$

and for the Rayleigh we have

$$Q_r = \sigma^2 \begin{bmatrix} 1 & 0 & \rho & \cdots & \rho^{n-1} \\ & & \ddots & & \vdots \\ & & & \ddots & \rho \\ & & & \ddots & 0 \\ & & & & 1 \end{bmatrix}$$

$$Q_r^{-1} = \frac{1}{\sigma^2(1-\rho)} \begin{bmatrix} 1 & 0 & -\rho & & & & \\ 0 & 1 & 0 & \ddots & & & \\ -\rho & 0 & 1+\rho^2 & \ddots & & & \\ \ddots & \ddots & \ddots & \ddots & & & \\ & & & & 1+\rho^2 & 0 & -\rho \\ & & & & 0 & 1 & 0 \\ & & & & -\rho & 0 & 1 \end{bmatrix}$$

$$|Q_r| = \sigma^{4n}(1-\rho^2)^{2(n-1)} \ .$$

The distribution for jointly Gaussian random variables is given by

$$\mathbf{P_0(x)} = \frac{1}{(2\pi)^{\frac{n}{2}} |Q|^{\frac{1}{2}}} \exp^{\frac{-1}{2}[(\mathbf{x}-\underline{\mu})^T Q^{-1}(\mathbf{x}-\underline{\mu})]} \ .$$

Lognormal:

Now in the lognormal case we have

$$\mathbf{x} = \begin{bmatrix} [x_{11}] & [x_{21}] & \cdots & [x_{n1}] \end{bmatrix}$$

and the 1 to 1 transformation $y_i = s \exp(x_{i1})$. Applying the Jacobian change of variables formula we get

$$|\mathbf{J}| = \frac{1}{\prod_{i=1}^n y_i}$$

so

$$
\begin{aligned}
P_Y(y_i; i=1,...,n) &= P_0(\ln(\frac{y_i}{s}); i=1,...,n) |\mathbf{J}| \\
&= \frac{1}{(2\pi)^{(n/2)} \sigma^n (1-\rho^2)^{\frac{(n-1)}{2}} \prod_{i=1}^n y_i} \\
&\quad \exp\left[\frac{-1}{2} \sum_{j=1}^n \sum_{k=1}^n q_{jk} \left[\ln(\frac{y_j}{s}) - \mu_j\right] \left[\ln(\frac{y_k}{s}) - \mu_k\right] \right]
\end{aligned}
$$

where q_{jk} is the jk^{th} element of Q_i^{-1} which can be rewritten as

$$P_Y(y_i; i = 1, ..., n) = \frac{1}{(2\pi)^{(n/2)}\sigma^n(1 - \rho^2)^{\frac{(n-1)}{2}} \prod_{i=1}^n y_i}$$

$$\exp\left[\frac{-1}{2\sigma^2(1 - \rho^2)}\left[\sum_{i=1}^n [i]^2 - 2\rho\sum_{i=1}^{n-1}[i][i+1] + \rho^2\sum_{i=2}^{n-1}[i]^2\right]\right]$$

where

$$[i] = \ln\left(\frac{y_i}{s}\right) - \mu$$

$$= \ln(y_i) - (\ln(s) + \mu).$$

Rayleigh:

Now for the Rayleigh case we have

$$\left[\begin{array}{cccc} [x_{11} & x_{12}] & [x_{21} & x_{22}] & \cdots & [x_{n1} & x_{n2}] \end{array}\right]$$

and the transformation $y_i = c\sqrt{x_{i1}^2 + x_{i2}^2}$. Then we have

$$P_Y(Y_i < y_i; i = 1, ..., n) = \int_{c\sqrt{x_{i1}^2 + x_{i2}^2} < y_i; i=1,...,n} P_0(\mathbf{x})d\mathbf{x}$$

introducing the variable θ and using the transformation

$$x_{i1} = y_i\cos(\theta_i)$$

$$x_{i2} = y_i\sin(\theta_i)$$

Then the Jacobian is given by

$$\mathbf{J} = \prod_{i=1}^n \left(\frac{y_i}{c^2}\right)$$

Defining the vector $\underline{v}$ as

$$\left[\begin{array}{cccc} [\frac{y_1}{c}\cos(\theta_1) & \frac{y_1}{c}\sin(\theta_1)] & [\frac{y_2}{c}\cos(\theta_2) & \frac{y_2}{c}\sin(\theta_2)] & \cdots & [\frac{y_n}{c}\cos(\theta_n) & \frac{y_n}{c}\sin(\theta_n)] \end{array}\right]$$

then

$$P_Y(Y_i < y_i; i = 1, ..., n) = \int_{\substack{Y_i < y_i, 0 < \theta < 2\pi \\ i=1,...,n}} P_0(\underline{v})\prod_{i=1}^n \left(\frac{Y_i}{c^2}\right)dY_id\theta_i$$

70

Assuming $\underline{\mu} = \underline{0}$, expanding the exponent and using the identity $\cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2) = \cos(\theta_1 - \theta_2)$ we get

$$
P_Y(Y_i < y_i; i = 1, ..., n) = \int_{\substack{Y_i < y_i; 0 < \theta < 2\pi \\ i=1,...,n}} \frac{\prod_{i=1}^n \frac{Y_i}{c^2}}{(2\pi)^{(n/2)} \sigma^{2n}(1 - \rho^2)^{(n-1)}}
$$
$$
\exp\left[\sum_{i=1}^n \alpha_i Y_i^2 \sum_{i=1}^{n-1} \beta Y_i Y_{i+1} \cos(\theta_i - \theta_{i+1})\right] \prod_{i=1}^n dY_i d\theta_i
$$

where

$$
\alpha_i = \begin{cases} \frac{-(1+\rho^2)}{2\sigma^2 c^2(1-\rho^2)} & i = 2, ..., n-1 \\ \frac{-1}{2\sigma^2 c^2(1-\rho^2)} & i = 1, n \end{cases}
$$

$$
\beta = \frac{\rho}{\sigma^2 c^2(1 - \rho^2)}
$$

Using the identity

$$
\int_{0 < \theta < 2\pi} \frac{1}{2\pi} \exp\left[A\cos(\theta - \theta_0)\right] d\theta = I_0(A)
$$

I_0 the modified zero order Bessel function. Then we get

$$
P_Y(Y_i < y_i; i = 1, ..., n) = \int_{Y_i < y_i; i=1,...,n} \frac{\prod_{i=1}^n Y_i}{\sigma^{2n} c^{2n}(1 - \rho^2)^{(n-1)}}
$$
$$
\exp\left[\sum_{i=1}^n \alpha_i Y_i^2\right] \prod_{i=1}^{n-1} I_0\left[\beta Y_i Y_{i+1}\right] \prod_{i=1}^n dY_i
$$

and so

$$
P_Y(y_i; i = 1, ...n) = \frac{\prod_{i=1}^n y_i}{\sigma^{2n} c^{2n}(1 - \rho^2)^{(n-1)}}
$$
$$
\exp\left[\sum_{i=1}^n \alpha_i y_i^2\right] \prod_{i=1}^{n-1} I_0\left[\beta y_i y_{i+1}\right] \quad .
$$

Appendix B

Mean, median, power, and variance of marginal lognormal and Rayleigh densities.

Lognormal:

$$y = s\exp(x) \; ; \; x \sim N(\mu, \sigma^2)$$

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma y} \exp\left\{ -\frac{\left[ln(\frac{y}{s}) - \mu \right]^2}{2\sigma^2} \right\} \; ; \; y\epsilon[0, \infty]$$

mean

$$
\begin{aligned}
\mathbf{E}[y] &= \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{\left[ln(\frac{y}{s}) - \mu \right]^2}{2\sigma^2} \right\} dy \\
&= s\exp\left\{ \mu + \frac{\sigma^2}{2} \right\}
\end{aligned}
$$

median

$$\int_0^r \frac{1}{\sqrt{2\pi}\sigma y} \exp\left\{ -\frac{\left[ln(\frac{y}{s}) - \mu \right]^2}{2\sigma^2} \right\} dy = \int_{-\infty}^{ln(\frac{r}{s})} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{[v - \mu]^2}{2\sigma^2} \right\} dv = \frac{1}{2}$$

The second integral looks like a Gaussian distribution and the median occurs at the mean so $ln\left(\frac{r}{s}\right) = \mu$

$$\text{median} = s\exp\{\mu\}$$

power

$$
\begin{aligned}
\mathbf{E}[y^2] &= \int_0^\infty \frac{y}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{\left[ln(\frac{y}{s}) - \mu \right]^2}{2\sigma^2} \right\} dy \\
&= s^2 \exp\left\{ 2(\mu + \sigma^2) \right\}
\end{aligned}
$$

variance

$$\mathbf{E}\left[(y - \mathbf{E}[y])^2\right] = \mathbf{E}[y^2] - \mathbf{E}[y]^2$$
$$= s^2 \exp\left\{2\mu + \sigma^2\right\}\left[\exp\{\sigma^2\} - 1\right]$$

Rayleigh:

$$y = c\sqrt{x_1^2 + x_2^2} \; ; \; x_i \text{ iid } N(0, \sigma^2)$$

$$f(y) = \frac{y}{\sigma^2 c^2}\exp\left\{-\frac{y^2}{2\sigma^2 c^2}\right\} \; ; \; y\epsilon[0,\infty]$$

mean

$$\mathbf{E}[y] = \int_0^\infty \frac{y^2}{\sigma^2}\exp\left\{-\frac{y^2}{2\sigma^2}\right\}dy$$
$$= \sqrt{\frac{\pi}{2}}c\sigma$$

median

$$\int_0^r \frac{y}{c^2\sigma^2}\exp\left\{-\frac{y^2}{2c^2\sigma^2}\right\}dy = 1 - \exp\left\{-\frac{r^2}{2c^2\sigma^2}\right\} = \frac{1}{2}$$
$$\text{median} = \sqrt{\ln(4)}c\sigma$$

power

$$\mathbf{E}[y^2] = \int_0^\infty \frac{y^3}{\sigma^2}\exp\left\{-\frac{y^2}{2\sigma^2}\right\}dy$$
$$= 2c^2\sigma^2$$

variance

$$\mathbf{E}\left[(y - \mathbf{E}[y])^2\right] = \mathbf{E}[y^2] - \mathbf{E}[y]^2$$
$$= \frac{(4 - \pi)c^2\sigma^2}{2}$$

Appendix C

Determination of stationary distribution of the underlying state equation

$$d\mathbf{x} = A\mathbf{x}dt + BdW$$

where A and B are constants. After discretizing we have

$$\mathbf{x}_{n+1} = (1 + A\Delta t)\mathbf{x}_n + B\Delta W_n \; ; \; \Delta W_n \text{ iid } N(0, \Delta t)$$

Denote $\hat{A} = 1 + A\Delta t$ then

$$\mathbf{x}_n = \hat{A}^n \mathbf{x}_0 + \sum_{i=0}^{n-1} \hat{A}^{n-1-i} B\Delta W_i \; ; \; \mathbf{x}_0 \sim N(0, \sigma^2)$$

Considering the scalar case which easily extends to the vector case for diagonal matrices we have

$$\mathbf{E}[\mathbf{x}_n] = 0$$

Denoting

$$y_i = \hat{A}^{n-1-i} B\Delta W_i$$

then

$$y_i \sim N(0, \hat{A}^{2(n-1-i)} B^2 \Delta t)$$

and

$$\hat{A}^n x_0 \sim N(0, \hat{A}^{2n} \sigma^2)$$

so

$$x_n \sim N(0, \hat{A}^{2n} \sigma^2 + \hat{A}^{2(n-1-i)} B^2 \Delta t)$$

Now. for a stationary distribution to exist $\hat{A}^{2n}\sigma^2 + \hat{A}^{2(n-1-i)}B^2\Delta t$ must converge. So $-1 < \hat{A} < 1$ or $\frac{-2}{\Delta t} < A < 0$. If we wish x_0 to be initialized to the stationary distribution then

$$\hat{A}^{2n}\sigma^2 + \sum_{i=0}^{n-1} \hat{A}^{2(n-1-i)}B^2\Delta t \overset{n\to\infty}{\Longrightarrow} \sigma^2$$

Assuming equality in the limit then

$$\frac{B^2\Delta t}{1-\hat{A}^{2n}} \sum_{i=0}^{n-1} \hat{A}^{2(n-1-i)} = \sigma^2$$

or

$$\frac{B^2\Delta t}{1-\hat{A}^{2n}} \sum_{m=0}^{n-1} \hat{A}^{2m} = \sigma^2$$

With the prior restriction $-1 < \hat{A} < 1$ the sum converges to $\frac{1}{1-\hat{A}^2}$ and $\hat{A}^{2n}$ converges to zero, so

$$\sigma^2 = \frac{B^2\Delta t}{1-\hat{A}^2} = \frac{B^2\Delta t}{1-(1+A\Delta t)^2}$$

So the diffusion equation has a stationary distribution of $N(0, \frac{B^2\Delta t}{1-(1+A\Delta t)^2})$ provided $\frac{-2}{\Delta t} < A < 0$.

Appendix D

We relate the underlying stationary distribution to underlying state diffusion parameters for the stochastic differential equation

$$d\mathbf{x} = A\mathbf{x}dt + BdW$$

where A and B are constants. After discretizing we have

$$\mathbf{x}_{n+1} = (1 + A\Delta t)\mathbf{x}_n + B\Delta W_n \; ; \; \Delta W_n \text{ iid } N(0, \Delta t) \; .$$

Denoting $\hat{A} = 1 + A\Delta t$ then

$$\mathbf{x}_n = \hat{A}^n \mathbf{x}_0 + \sum_{i=0}^{n-1} \hat{A}^{n-1-i} B\Delta W_i \; ; \; \mathbf{x}_0 \sim N(0, \sigma^2) \; .$$

From Appendic C we have $\mathbf{E}[\mathbf{x}_n] = 0$ so for the scalar case

$$
\begin{aligned}
R(k) &= \mathbf{E}[\mathbf{x}_{n+k}\mathbf{x}_n] = \hat{A}^k \mathbf{x}_n^2 + \sum_{j=0}^{k-1} \hat{A}^j B\Delta W_{n+k-1-j}\,\mathbf{x}_n \\
&= \hat{A}^k \mathbf{E}[\mathbf{x}_n^2] \\
&= \hat{A}^k \sigma^2 \; .
\end{aligned}
$$

If we set $\hat{A} = \exp\left\{-\frac{\Delta t}{t_c}\right\}$ then x_n has exponential decorrelation with time constant t_c and

$$\rho = \exp\left\{-\frac{\Delta t}{t_c}\right\} = \hat{A} = 1 - A\Delta t$$

or

$$A = \frac{\exp\left\{-\frac{\Delta t}{t_c}\right\} - 1}{\Delta t}$$

and from Appendix C

$$B = \sigma \sqrt{\frac{1 - (1 + A\Delta t)^2}{\Delta t}} \quad .$$

Appendix E

The following tables summarize the simulation results predominantly discussed in Chapter 4. Where unspecified $P_F = P_M = .001$.

| | $\approx t_r$ | Hypothesis 0 | | | | |
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	63.4	190.7	36.6	1341.3	611.4
means	$10t_l$	31.5	327.0	68.5	1129.6	877.2
match	t_l	94.9	761.1	5.1	994.1	773.0
	$t_l/10$	100.0	114.2	0	-	114.2
	$t_l/100$	100.0	29.9	0	-	29.9
	$100t_l$	74.9	284.8	25.1	1341.8	549.9
	$10t_l$	36.8	430.8	63.2	1301.4	981.4
	t_l	96.8	781.2	3.2	1099.6	791.5
	$t_l/10$	100.0	123.9	0	-	123.9
	$t_l/100$	100.0	30.2	0	-	30.2

| | $\approx t_r$ | Hypothesis 1 | | | | |
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	100.0	700.4	0	-	700.4
means	$10t_l$	100.0	780.9	0	-	780.9
match	t_l	100.0	936.9	0	-	936.9
	$t_l/10$	96.4	1317.2	3.6	788.6	1298.4
	$t_l/100$	0	-	100.0	48.8	48.8
	$100t_l$	100.0	776.4	0	-	776.4
	$10t_l$	100.0	915.0	0	-	915.0
	t_l	100.0	1091.2	0	-	1091.2
	$t_l/10$	93.7	1616.4	6.3	1241.2	1592.7
	$t_l/100$	0	-	100.0	47.8	47.8

Table E.1: Implicit scheme, SNR=2 $P_M = P_F = .001$

| | $\approx t_r$ | Hypothesis 0 | | | | |
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	63.4	190.7	36.6	1341.7	611.6
means	$10t_l$	31.4	320.3	68.6	1135.6	880.0
match	t_l	95.5	762.8	5.5	981.8	774.8
	$t_l/10$	100.0	113.0	0	-	113.0
	$t_l/100$	100.0	29.1	0	-	29.1
	$100t_l$	79.4	284.7	25.1	1342.3	549.9
	$10t_l$	37.2	434.7	62.8	1302.5	979.7
	t_l	96.7	800.3	3.3	1088.2	809.8
	$t_l/10$	100.0	122.4	0	-	122.4
	$t_l/100$	100.0	29.4	0	-	29.4

| | $\approx t_r$ | Hypothesis 1 | | | | |
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	100.0	699.3	0	-	699.3
means	$10t_l$	100.0	781.9	0	-	781.9
match	t_l	100.0	932.1	0	-	932.1
	$t_l/10$	96.3	1263.6	3.7	758.7	1244.7
	$t_l/100$	0	-	100.0	48.3	48.3
	$100t_l$	100.0	775.0	0	-	775.0
	$10t_l$	100.0	915.7	0	-	915.7
	t_l	100.0	1085.0	0	-	1085.0
	$t_l/10$	94.0	1574.6	^.0	1185.9	1551.4
	$t_l/100$	0	-	100.0	47.0	47.0

Table E.2: Mixed scheme, SNR=2 $P_M = P_F = .001$

| | $\approx t_r$ | Hypothesis 0 | | | | |
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	63.0	72.7	37.0	658.6	289.7
means	$10t_l$	25.5	113.3	74.5	576.6	458.5
match	t_l	79.5	487.2	20.5	664.7	523.5
	$t_l/10$	100.0	68.2	0	-	68.2
	$t_l/100$	100.0	22.4	0	-	22.4
	$100t_l$	70.1	75.8	29.9	824.3	299.3
	$10t_l$	35.6	133.9	64.4	637.0	457.9
	t_l	83.3	493.7	16.7	741.4	535.1
	$t_l/10$	100.0	72.2	0	-	72.0
	$t_l/100$	100.0	21.4	0	-	21.4

| | $\approx t_r$ | Hypothesis 1 | | | | |
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	100.0	419.7	0	-	419.7
means	$10t_l$	100.0	468.9	0	-	468.9
match	t_l	100.0	500.6	0	-	500.6
	$t_l/10$	99.5	393.0	0.5	164.4	391.9
	$t_l/100$	0	-	100.0	78.5	78.5
	$100t_l$	100.0	488.5	0	-	488.5
	$10t_l$	100.0	508.0	0	-	508.0
	t_l	100.0	597.0	0	-	597.0
	$t_l/10$	99.8	456.1	0.2	400.0	456.0
	$t_l/100$	0	-	100.0	73.9	73.9

Table E.3: Implicit Scheme, SNR=5 $P_M = P_F = .001$

	$\approx t_r$	Hypothesis 0				
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	63.0	72.7	37.0	658.6	289.7
means	$10t_l$	25.8	117.0	74.2	576.4	457.9
match	t_l	79.3	498.5	20.7	681.2	536.4
	$t_l/10$	100.0	67.0	0	-	67.0
	$t_l/100$	100.0	22.1	0	-	22.1
	$100t_l$	70.1	75.8	29.9	824.3	299.3
	$10t_l$	35.2	135.5	64.3	637.6	461.0
	t_l	83.9	502.6	16.1	731.8	539.4
	$t_l/10$	100.0	71.2	0	-	71.2
	$t_l/100$	100.0	21.0	0	-	21.0

	$\approx t_r$	Hypothesis 1				
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	100.0	419.3	0	-	419.3
means	$10t_l$	100.0	468.6	0	-	468.6
match	t_l	100.0	499.4	0	-	499.4
	$t_l/10$	99.6	380.0	0.4	153.0	380.0
	$t_l/100$	0	-	100.0	84.7	84.7
	$100t_l$	100.0	489.1	0	-	489.1
	$10t_l$	100.0	508.1	0	-	508.1
	t_l	99.9	595.1	0.1	341.0	594.9
	$t_l/10$	99.9	439.5	0.1	309.0	439.4
	$t_l/100$	0	-	100.0	79.2	79.2

Table E.1: Mixed scheme, SNR=5 $P_M = P_F = .001$

	$\approx t_r$	Hypothesis 0				
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	72.7	16.3	27.3	219.2	71.7
means	$10t_l$	31.8	22.7	68.2	209.5	150.1
match	t_l	40.6	94.6	59.4	207.0	161.3
	$t_l/10$	99.1	28.4	0.9	52.0	28.6
	$t_l/100$	100.0	10.1	0	-	10.1
	$100t_l$	91.3	15.0	8.7	431.0	51.2
	$10t_l$	44.6	17.3	55.4	240.9	141.1
	t_l	50.0	71.7	50.0	238.0	154.8
	$t_l/10$	99.5	36.3	0.5	43.5	36.4
	$t_l/100$	100.0	10.9	0	-	10.9

	$\approx t_r$	Hypothesis 1				
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	99.8	180.6	0.2	584.0	180.4
means	$10t_l$	100.0	199.9	0	-	199.9
match	t_l	100.0	233.5	0	-	233.5
	$t_l/10$	100.0	144.9	0	-	144.9
	$t_l/100$	92.5	132.6	7.5	85.0	129.1
	$100t_l$	100.0	206.7	0	-	206.7
	$10t_l$	100.0	224.7	0	-	224.7
	t_l	100.0	237.5	0	-	237.5
	$t_l/10$	100.0	181.0	0	-	181.0
	$t_l/100$	91.7	224.8	8.3	248.0	226.8

Table E.2: Implicit scheme, SNR=20 $P_M = P_F = .001$

	$\approx t_r$	Hypothesis 0				
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	72.7	16.3	27.3	219.2	71.7
means	$10t_l$	32.1	25.6	67.9	217.3	155.8
match	t_l	41.3	93.9	58.7	204.2	158.7
	$t_l/10$	99.1	27.9	0.9	52.0	28.1
	$t_l/100$	100.0	9.9	0	-	9.9
	$100t_l$	99.1	14.7	8.1	518.8	55.5
	$10t_l$	44.6	17.3	55.4	240.9	141.1
	t_l	50.0	72.0	50.0	237.8	154.9
	$t_l/10$	99.5	35.9	0.5	43.5	35.9
	$t_l/100$	100.0	10.8	0	-	10.8

	$\approx t_r$	Hypothesis 1				
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
stat.	$100t_l$	99.8	180.5	0.2	584.0	181.3
means	$10t_l$	100.0	203.3	0	-	203.3
match	t_l	100.0	233.3	0	-	233.3
	$t_l/10$	100.0	136.9	0	-	136.9
	$t_l/100$	95.5	87.6	4.5	56.7	86.2
	$100t_l$	100.0	202.0	0	-	202.7
	$10t_l$	100.0	224.7	0	-	224.7
	t_l	100.0	274.2	0	-	274.2
	$t_l/10$	100.0	175.0	0	-	175.0
	$t_l/100$	98.6	118.7	1.4	37.0	117.5

Table E.3: Mixed scheme, SNR=20 $P_M = P_F = .001$

		Hypothesis 0				
P_f	P_m	% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
.001	.0001	100.0	85.5	0	-	85.5
.001	.001	100.0	68.2	0	-	68.2
.001	.01	100.0	49.4	0	-	49.4
.001	.1	100.0	26.3	0	-	26.3
.001	.5	100.0	14.9	0	-	14.9
.001	.75	100.0	9.8	0	-	9.8

		Hypothesis 1				
P_f	P_m	% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
.001	.0001	100.0	392.9	0	-	392.9
.001	.001	99.5	393.0	.5	164.4	391.9
.001	.01	98.0	371.6	2.0	137.5	367.0
.001	.1	88.8	357.8	11.2	101.0	329.0
.001	.5	53.1	281.1	46.9	34.7	165.6
.001	.75	41.0	226.1	59.0	22.1	105.7

Table E.7: Implicit scheme, SNR=5, varying the probability of a miss

P_f	P_m	Hypothesis 0				
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
.0001	.001	100.0	69.9	0	-	69.9
.001	.001	100.0	68.2	0	-	68.2
.01	.001	100.0	66.1	0	-	66.1
.1	.001	98.1	66.0	1.9	51.4	65.7
.5	.001	63.9	47.8	36.1	10.8	34.5
.75	.001	36.4	38.8	63.6	3.6	16.4

P_f	P_m	Hypothesis 1				
		% correct	avg. det. time correct	% wrong	avg. det. time wrong	avg. det. time
.0001	.001	99.8	506.2	.2	332.0	505.8
.001	.001	99.5	393.0	.5	164.4	391.9
.01	.001	99.7	264.4	.3	109.3	263.9
.1	.001	99.9	129.4	.1	222.0	129.5
.5	.001	99.9	26.2	.1	370.0	26.5
.75	.001	99.8	6.8	.2	86.5	7.0

Table E.8: Implicit scheme, SNR=5, varying the probability of false alarm

Bibliography

[Bar64]   David K. Barton. *Radar System Analysis*. Artech, 1964.

[Bar80]   John S. Baras. *Ship RCS Scintillation Simulation*. Unclassified Editted Technical Report of NRL Report 8189, Naval Research Lab. 1980.

[BBH83]   J.S. Baras, G.L. Blankenship, and W.E. Hopkins, Jr. Existence, Uniqueness, and Asymptotic Behavoir of Solution to a Class of Zakai Equations with Unbounded Coefficients. *IEEE Trans. Auto Control*, 203–214, 1983.

[Bla86]   Lamont V. Blake. *Radar Range Performance Analysis*. Artech House, Inc., 1986.

[Don79]   J. J. Dongarra, et al. *Linpack: Users' Guide*. Philadelphia: Society for Industrial and Applied Mathematics, 1979.

[ER87]   Jerry L. Eaves and Edward K. Reedy. *Priciples of Modern Radar*. Van Nostrand Reinhold Co., 1987.

[GL89]   Evaggelos Geraniotis and Joseph Lawrence. *Chaff-Target Discrimination from Correlated Envelope Observations, Part I :Optimal and Near Optimal Block Discriminators*. Technical Report, Naval Research Lab, 1989.

[Hub81]   P. J. Huber. *Robust Statistics*. Wiley-Interscience, 1981.

[LaV86] Anthony LaVigna. *Real Time Sequential Detection for Diffusion Signals*. Master's thesis, University of Maryland, College Park MD, 1986.

[LS77]   R. S. Liptser and A. N. Shiryayev. *Statistics of Random Processes I General Theory*. Springer-Verlag, 1977.

[LS78]   R. S. Liptser and A. N. Shiryayev. *Statistics of Random Processes II Applications*. Springer-Verlag, 1978.

[MAC86] MACSYMA Group of Symbolics, Inc. *Zetalisp MACSYMA Reference Manual*. Symbolic, Inc., 1986.

[MM73]   Daniel P. Meyer and Herbert A. Mayer. *Radar Target Detection Handbook of Theory and Practice*. Academic Press, 1973.

[Sko80] Merril I. Skolnik. *Introduction to Radar Systems*. McGraw Hill Bool Co., 1980.

[VK81]   V. Vemuri and Walter Karplus. *Digital Computer Treatment of Partial Differential Equations*. Prentice-Hall, 1981.

[WH85]   E. Wong and B. Hajek. *Stochastic Processes in Engineering Systems*. Springer-Verlag, 1985.